On branching process with rare neutral mutation[*]

Airam Blancas Benítez [†]   Víctor Rivero [†]

Abstract

In this paper we study the genealogical structure of a Galton-Watson process with neutral mutations, where the initial population is large and mutation rate is small [3]. Namely, we extend in two directions the results obtained in Bertoin's work. In the critical case, we construct the version of Bertoin's model conditioned not to be extinct, and in the case with finite variance we show convergence after normalization, of allelic sub-populations towards a tree indexed CSBP with immigration. Besides, we establish the version of the limit theorems in [3], been for the unconditioned process and for the process conditioned to non-extinction, in the case where the reproduction law has infinite variance and it is in the domain of attraction of an α-stable distribution.

Keywords. Branching process; Neutral mutations; Q-processes; Regular variation; Domain of attraction of α-stable laws.

1   Introduction models and main results

A Galton-Watson process models a population where at every generation each individual reproduces according to the same distribution, independently of the others and then dies. A number of variants, involving different types of conditioning and limit theorems, are core of branching processes theory. For instance, when the process dies with probability one, Yaglom (1947) proved that the distribution of the process conditioned to non-extinction exists, under some assumptions on the moments of the reproduction law. The proof was simplified and the moment assumptions removed by Joffe (1967), Seneta and Vere-Jones (1967). More generally, Lamperti and Ney (1968) introduced the Q-process, conditioning on not being extinct in the distant future and being extinct in the even more distant future. For background about branching processes we refer to [1], [12] and [13].

As an extension of the Galton-Watson model, Bertoin [3] studied the so called Galton-Watson process with neutral mutations; this emerges assuming that the mutations modify the genotype of individuals but not the dynamic of a population modeled by a standard Galton-Watson. Since mutations appear in the ancestral lines of the population, each individual begets children that do not necessarily inherit its genetic type (allele). In addition, we suppose that the population

has infinity alleles, that is, each mutation event originates a different allele. We denote the size of a typical family size by $\xi^{(+)} := \xi^{(c)} + \xi^{(m)}$, where $\xi^{(c)}$, $\xi^{(m)}$ are non-negative random variables which determine the number of clones and mutants children of a typical individual, respectively. We exclude the degenerate cases $\xi^{(c)} \equiv 0$ or $\xi^{(m)} \equiv 0$.

In [3], Bertoin established asymptotic features on the genealogy of allelic sub-families in a Galton-Watson process with neutral mutations. In his development, the genealogy of the population is described by a planar rooted tree where the mutations are represented by marks in the edges between parents and mutant childrens. The vertices with n marks in their ancestral line are associated with the called n-type individuals; those individuals with parent of $(n-1)$-type are known mutants of the n-type. We denote by T_n the total population of individuals of the n-th type and by M_n the total number of mutants of n-th type, with the convention that mutants of the 0-th type are the ancestors, that is $M_0 = a$, $\mathbb{P}_a$-c.s. Observe that for every n each mutant of the n-type begets a sub-family, which is independent of the others and has the same behavior that the original tree. In fact, this is a consequence of the general branching property, which states that conditionally on the set of children of a stopping line, the families that those beget are independent copies of the initial tree. The concept of stopping line was introduced formally by Chauvin [5]. Namely, a line is a family of edges such that every branch from the root contains at most one edge in that family. One may think that a stopping line is a random line such that the event, "an edge is in the line", only depends on the marks found on their ancestral line. An important consequence from the general branching property is the following lemma.

Lemma 1.1 (Bertoin [3], Lemma 1). *Under $\mathbb{P}_a$*

$$\{M_n : n \in \mathbb{Z}_+\},$$

is a Galton-Watson process with reproduction law $\mathbb{P}_1(M_1 \in \cdot)$. More generally,

$$\{(T_n, M_{n+1}) : n \in \mathbb{Z}_+\},$$

is a Markov chain, with transition probabilities

$$\mathbb{P}_a(T_n = k, M_{n+1} = l \mid T_{n-1} = i, M_n = j) = \mathbb{P}_j(T_0 = k, M_1 = l), \quad i, j, k, l \in \mathbb{Z}_+ \text{ and } j \le k. \quad (1.1)$$

Remark 1.2. Since the mutants of the n-th type are also individuals of the n-th type, the transition probability in (1.1) is zero when $j > k$.

For notational convenience $\{P^n_{(i,j),(k,l)} : i, j, k, l \in \mathbb{Z}_+\}$ denotes the n-step transition probabilities of $\{(T_n, M_{n+1}), n \in \mathbb{Z}_+\}$, that is

$$P^n_{(i,j),(k,l)} = \mathbb{P}_a(T_{r+n} = k, M_{r+n+1} = l | T_r = i, M_{r+1} = j), \quad i, j, k, l \in \mathbb{Z}_+, n \in \mathbb{N}. \quad (1.2)$$

Remark 1.3. For $i, j, k, l \in \mathbb{Z}_+$, $P^n_{(i,j),(k,l)}$ depends only on the second coordinate, actually is not difficult to prove using induction that

$$P^n_{(i,j),(k,l)} = \sum_{j_{n-1}=1}^{\infty} \mathbf{P}^{n-1}_{(j,j_{n-1})} \mathbb{P}_{j_{n-1}}(T_0 = k, M_1 = l), \quad (1.3)$$

where $j_0 = j$ and $\{\mathbf{P}^n_{(i,j)} : i, j \in \mathbb{Z}_+\}$ denotes the n-step transition probabilities of $\{M_n, n \in \mathbb{Z}_+\}$.

Let $\mathbb{U}$ be the finite sequence of integers

$$\mathbb{U} := \bigcup_{k \in \mathbb{Z}_+} \mathbb{N}^k,$$

where $\mathbb{N} = \{1, 2, ...\}$ and $\mathbb{N}^0 = \{\varnothing\}$. We consider a locally finite rooted tree $\mathcal{A} = \{\mathcal{A}_u : u \in \mathbb{U}\}$ called tree of alleles. According to Ulam-Harris-Neveu labeling the root is $\varnothing$, a vertex at its level $k > 0$ is $u = (u_1, ..., u_k)$ and $uj = (u_1, ..., u_k, j)$ is the j-th children. We denote by $|u|$ the level of the vertex u. A tree $\mathcal{A}$, is constructed recursively in [3] taking $\mathcal{A}_\varnothing = T_0$ and $\mathcal{A}_{uj}$ as the size of the j-th allelic sub-population of the type $|u| + 1$ which descend from the allelic sub-family indexed by the vertex u, with the convention that in the case of ties, sub-families are ordered uniformly at random. Since the transition probabilities of the chain $\{(T_n, M_{n+1}) : n \in \mathbb{Z}\}$ depend only on the second coordinate, the tree of alleles enjoy a kind of branching property, the formal statement is given in the following result, where the notation $(d_u \downarrow)$ means that the d_u-tuple has been rearranged in the decreasing order of the first coordinate, by convention, in the case of ties, the coordinates are ranked uniformly at random.

Lemma 1.4 (Bertoin [3], Lemma 2)**.** *For every integer $a \geq 1$ and $k \geq 0$, under $\mathbb{P}_a$ conditionally on $((\mathcal{A}_u, d_u) : |u| \leq k)$, for each vertex u at level k with $\mathcal{A}_u > 0$, the family of variables $((\mathcal{A}_{uj}, d_{uj}) : 1 \leq j \leq d_u)$ are independent with distribution $(T_0, M_1)^{(d_u \downarrow)}$ under $\mathbb{P}_1$.*

It is important to observe that if $d_u := \max\{j \geq 1 : \mathcal{A}_{uj} > 0\}$, agreeing that $\max \varnothing = 0$, then

$$T_k = \sum_{|u|=k} \mathcal{A}_u \quad \text{and} \quad M_{k+1} = \sum_{|u|=k} d_u. \tag{1.4}$$

Hence given a population with neutral mutations, the tree of alleles records the genealogy of allelic sub-families together with their sizes. Also, the size of their generations is a Galton-Watson process.

Besides, we consider for every integer n, a Galton-Watson process started from n ancestors, with a fixed reproduction law, which is critical with finite variance and mutations occur at random with rate $1/n$. The main result of [3] states that as n tends to infinity, the rescaled tree of alleles $n^{-2}\mathcal{A}^{(n)}$ converges in the sense of finite dimensional distributions towards a process $\{\mathcal{Z}_u : u \in \mathbb{U}\}$, called tree-indexed CSBP with values in $(0, \infty)$, such that the vertices on a level are distributed as the family of the atoms of a Poisson random measure conditionally on the preceding level. This result gives another point of view to the convergence of a Galton-Watson process rescaled towards a discrete time continuous state branching process, CSBP for short.

A first goal in this paper is to construct the chain $\{(T_n, M_{n+1}) : n \in \mathbb{Z}_+\}$ conditioned on non-extinction of the mutants, so we are interested in the situation where

$$T = \inf\{n \geq 1 : M_n = 0\} < \infty, \quad \mathbb{P}_a\text{-}c.s. \tag{1.5}$$

That is the purpose of the following Theorem.

Theorem 1.5. *Let $a \in \mathbb{Z}_+$ and $\mathcal{F}_n$ the natural filtration of the process $\{(T_{n-1}, M_n) : n \in \mathbb{N}\}$. There exits a probability measure $\mathbb{P}_a^\uparrow$ that can be expressed as a h-transform of $\mathbb{P}_a$ using the $(\mathcal{F}_n)$-martingale*

$$Y_n = \frac{M_n q^{M_n - a}}{(f'(q))^n} \mathbf{1}_{\{n < T\}},$$

where $f(y) = \mathbb{E}_1(y^{M_1})$ and $q = \mathbb{P}(0 < T < \infty)$.

That is
$$d\mathbb{P}^\uparrow_a|_{\mathcal{F}_n} = \frac{Y_n}{a} d\mathbb{P}_a|_{\mathcal{F}_n}, \quad n \in \mathbb{N}.$$

Furthermore, $\mathbb{P}^\uparrow_a$ is the law of a Markov chain $\{(T^\uparrow_n, M^\uparrow_{n+1}), n \in \mathbb{Z}_+\}$ with n-step transition probabilities,

$$Q^n_{(i,j),(k,l)} = \frac{lq^{l-j}}{j(f'(q))^n} P^n_{(i,j),(k,l)}, \quad j, l \geq 1, \tag{1.6}$$

where $\{P^n_{(i,j),(k,l)} : i, j, k, l \in \mathbb{Z}_+\}$ denotes the n-step transition probabilities of $\{(T_n, M_{n+1}), n \in \mathbb{Z}_+\}$.

Next we ensure that the process defined in the above Theorem is distributed as $\{(T_n, M_{n+1}) : n \in \mathbb{Z}\}$ conditionally on non-extinction of mutants. Actually, thanks to a relation that we establish in Section 2, between the generating functions of (T_n, M_{n+1}) and M_n, we can use classical methods and prove the following result.

Theorem 1.6. *Suppose that $\mathbb{E}(\xi^{(c)}) < 1$ and $\mathbb{E}(\xi^{(+)}) \leq 1$.*

i) *The Yaglom limit*
$$\lim_{n \to \infty} \mathbb{P}(T_{n-1} = i, M_n = j \,|\, n < T < \infty),$$

exists and has a generating function $\widehat{\varphi}(x, y)$ such that for all $n \in \mathbb{N}$,

$$m^n \widehat{\varphi}(x, y) = \widehat{f}(\varphi_n(x, y)) - \widehat{f}(\varphi_n(x, 0)), \quad x, y \in [0, 1]. \tag{1.7}$$

ii) *Let $a \in \mathbb{Z}_+$ and n fixed. The conditional laws of the process $\{(T_k, M_{k+1}) : 0 \leq k \leq n-1\}$ under $\mathbb{P}_a(\cdot \,|\, n + k < T < \infty)$ converge towards the probability measure $\mathbb{P}^\uparrow_a$, in the sense that for any $n \geq 0$*
$$\lim_{k \to \infty} \mathbb{P}_a(A \,|\, n + k < T < \infty) = \mathbb{P}^\uparrow_a(A), \quad \forall A \in \mathcal{F}_n. \tag{1.8}$$

Remark 1.7. In the situation $\mathbb{E}(\xi^{(c)}) < 1$, according to Corollary 1 of [3], the hypothesis $\mathbb{E}(\xi^{(+)}) \leq 1$ is equivalent to the condition (1.5).

Furthermore, in Section 3.3 we will see that just as in the case of Galton-Watson processes conditioned to non-extinction, the dynamic of $\{(T^\uparrow_n, M^\uparrow_{n+1}) : n \in \mathbb{Z}\}$ admits a description in terms of immigration of mutants of the n-th type into the population.

Besides we investigate, in the same sense of [3] but on a complimentary class of reproduction laws, the asymptotic behavior of the population. Namely, we consider a Galton-Watson process $\{Z^+_n, n \in \mathbb{Z}_+\}$ such that the reproduction law

$$\pi^+_k = \mathbb{P}(\xi^{(+)} = k), \quad k \in \mathbb{Z}_+,$$

is critical but, instead of assumming that it has finite variance as in Bertoin's paper, we suppose that there exists $\alpha \in (1, 2)$ such that,

$$\bar{\pi}^+(j) := \mathbb{P}(\xi^{(+)} > j) \in RV^{-\alpha}_\infty, \tag{1.9}$$

where $RV_\infty^{-\alpha}$ denotes the class of functions which are regularly varying with index $-\alpha$ at ∞, see Chapter I in [4] for background. Note that the case where $\alpha \in (0, 1)$ is not allowed here because it contradicts the assumption π^+ is critical.

In order to extended the main result of [3] to our setting, we use that there exists a regularly varying function r with index $1/\alpha$ such that

$$r(n)\mathbb{P}(\xi^+ > ny) \xrightarrow[n\to\infty]{} c_\alpha y^{-\alpha}, \quad \forall y > 0, \tag{1.10}$$

where $c_\alpha = 1/\Gamma(3 - \alpha)$. A proof of this fact is given in Proposition 4.1. Also, we assume that each child is a clone of her mother with fixed probability $1 - p$ and a mutant with probability p, so the joint law of $(\xi^{(c)}, \xi^{(m)})$, denoted by $\pi = (\pi_{k,l} : k, l \in \mathbb{Z}_+)$,

$$\pi_{k,l} = \mathbb{P}(\xi^{(c)} = k, \xi^{(m)} = l), \quad k, l \geq 0, \tag{1.11}$$

satisfies

$$\pi_{k,l} = \pi_{k+l}^+ \binom{k + l}{k}(1 - p)^k p^l, \quad k, l \in \mathbb{Z}_+. \tag{1.12}$$

In the paper [3], it has been assumed that the number of ancestors and mutation rate have the following behavior

$$a(n) \sim nx \quad \text{and} \quad p(n) \sim cn^{-1}, \quad \text{as } n \to \infty. \tag{1.13}$$

where c, x are some positive constants. Here we will instead assume that following behaviour

$$a(n) \sim xr(n)p(n) \quad \text{and} \quad p(n) \sim cn^{-1}, \quad \text{as } n \to \infty. \tag{1.14}$$

In the remainder of this paper the relation $f \sim g$ refers to $\lim_{x\to\infty} f(x)/g(x) = 1$. The notation $\implies$ refers to convergence in distribution as $n \to \infty$ and $\mathcal{L}\left(\cdot, \mathbb{P}_{a(n)}^{p(n)}\right)$ to the distribution of the process under $\mathbb{P}_{a(n)}^{p(n)}$. The result below extends to our setting the main result in [3].

Theorem 1.8. *If (1.9) and (1.14) holds. Then, the following convergence holds in the sense of finite dimensional distributions*

$$\mathcal{L}\left((((r(n))^{-1}\mathcal{A}_u^{(n)}, (r(n)p(n))^{-1}d_u^{(n)}) : u \in \mathbb{U}), \mathbb{P}_{a(n)}^{p(n)}\right) \implies ((\mathcal{Z}_u^{1/\alpha}, \mathcal{Z}_u^{1/\alpha}) : u \in \mathbb{U}),$$

where $\{\mathcal{Z}_u : u \in \mathbb{U}\}$ is a tree-indexed CSBP with reproduction measure

$$\nu^\alpha(dy) = c_\alpha' x^{-1-1/\alpha}dy, \quad y > 0, \alpha \in (1, 2), \tag{1.15}$$

where $c_\alpha' = \alpha^{-1}/\Gamma(1 - \alpha^{-1})$.

Finally we establish the convergence of the finite dimensional distributions of the rescaled chain $\{T_n, M_{n+1}\}$, conditioned to non-extinction of mutants, towards a continuous state branching process with immigration in discrete time.

Theorem 1.9. *If the reproduction law is critical, there exist sequences $b_1(n)$ and $b_2(n)$ such that the following joint convergence in the sense of finite dimensional distributions hold:*

$$\mathcal{L}\left((b_1(n)T_{k-1}, b_2(n)M_k) : k \in \mathbb{Z}_+), \mathbb{P}_{a(n)}^{p(n)\uparrow}\right) \implies ((Y_k, \beta Y_k) : k \in \mathbb{Z}_+),$$

where $\{Y_k : k \in \mathbb{Z}_+\}$ is a CSBP with immigration such that

a) when the reproduction law has finite variance σ^2 and (1.13) holds, then its reproduction measure is

$$\nu(dy) = \frac{c}{\sqrt{2\pi\sigma^2 y^3}} \exp\left(-\frac{c^2 y}{2\sigma^2}\right) dy, \quad y > 0, \tag{1.16}$$

then the immigration measure is $z\nu(dz)$ and $\beta = c$; moreover $b_1(n) = n^{-2}$ and $b_2(n) = n^{-1}$;

b) when the assumptions (1.9) and (1.14) hold, the reproduction measure is $\nu^\alpha(dz)$ as defined in (1.15), the immigration measure is $z\nu^\alpha(dz)$ and $\beta = 1$; the normalizing constants are given by $b_1(n) = (r(n)p(n))^{-1}$ and $b_2(n) = (r(n))^{-1}$.

The remainder of the paper is organized as follows. In Section 2 we provide some preliminary facts. Section 3, is devoted to construct and to interpret the process $\{(T_n, M_{n+1}) : n \in \mathbb{Z}\}$ conditioned to non extinction. The last section is divided in three parts; the first one correspond to the framework where the reproduction law is in the domain of attraction of an α-stable distribution. Section 4.2 includes the proof of our main result in this setting. Finally we prove Theorem 1.9, which is the result that explains the asymptotic behaviour of the process conditioned to non-extinction.

2 Preliminaries

In this section we obtain some useful formulas for the generating function of (T_n, M_{n+1}), denoted for $n \in \mathbb{Z}$ by

$$\varphi_n(x, y) := \mathbb{E}_1(x^{T_{n-1}} y^{M_n}), \quad x, y \in [0, 1],$$

where for notational convenience $\varphi_1(x, y) := \varphi(x, y)$. Observe that the generating function of M_n is

$$f_n(y) := \varphi_n(1, y), \quad y \in [0, 1], \tag{2.1}$$

and as before we denote $f_1(y) =: f(y)$.

According to classical theory of branching processes, the extinction probability of the Galton-Watson process $\{M_n, n \in \mathbb{Z}_+\}$, that we denote by q, is the smallest root of $f(y) = y$, which is less or equal than one depending on whether, the mean of the reproduction law, $m := \mathbb{E}(M_1)$ is > 1, or ≤ 1, respectively. In order to avoid trivial cases, we assume throughout that

H1) $\mathbb{P}(M_1 = 1) > 0$,

H2) $\mathbb{P}(M_1 = 0) + \mathbb{P}(M_1 = 1) < 1$, and $\mathbb{P}(M_1 = j) \neq 1$, for any j.

We also know that the n-step transition probabilities $\{\mathbf{P}_{(i,j)}^n : i, j \in \mathbb{Z}_+\}$ of $\{M_n, n \in \mathbb{Z}_+\}$ satisfy

$$\sum_{j=0}^{\infty} \mathbf{P}_{(i,j)}^n y^j = (f_n(y))^i, \quad i \geq 1. \tag{2.2}$$

Now for a Galton-Watson process with neutral mutations, let g be the generating function of the reproduction law of a typical individual, that is

$$g(x, y) := \mathbb{E}(x^{\xi^{(c)}} y^{\xi^{(m)}}), \quad x, y \in [0, 1].$$

Proposition 1 of [3] ensure that the law of (T_0, M_1) can be obtained applying the Lagrange inversion formula to the equation

$$\varphi(x, y) = xg(\varphi(x, y), y), \quad x, y \in [0, 1], \tag{2.3}$$

which gives the identity

$$\mathbb{P}_a(T_0 = k, M_1 = l) = \frac{a}{k}\pi_{k-a,l}^{*k}, \quad k \geq a \geq 1 \text{ and } l \geq 0, \tag{2.4}$$

where π^{*k} denotes the k-th convolution of π, as defined in (1.11).

The statement in the remark 1.7 is a consequence of (2.3). We also have, $\mathbb{E}(\xi^{(+)2}) < \infty$ if and only if $\mathbb{E}(M_1^2) < \infty$. From (2.4) we can write the hypothesis (H1) and (H2) in terms of the reproduction distribution of a typical individual. Besides, we have a similar identity to (2.2) for the process $\{(T_n, M_{n+1}), n \in \mathbb{Z}_+\}$

$$\sum_{k,l=0}^{\infty} P_{(i,j),(k,l)}^n x^k y^l = (\varphi_n(x, y))^j, \quad i, j \geq 1, \tag{2.5}$$

where P^n denotes the n-step transition probabilities. We get this equality by induction, for we apply the Chapman-Kolmogorov equation to express the $(n + 1)$-step transition probabilities in terms of the transitions in one step and use (1.1). A simple but key relation for our analysis is

$$\varphi_n(x, y) = f_{n-1}(\varphi(x, y)), \quad x, y \in [0, 1]. \tag{2.6}$$

Due to (2.1), the proof of the identity is equivalent to establish the following equality

$$\varphi_n(x, y) = \varphi_{n-1}(1, \varphi(x, y)), \quad x, y \in [0, 1], \tag{2.7}$$

which holds from the following standard calculations

$$\begin{aligned}
\varphi_n(x, y) &= \mathbb{E}_1(\mathbb{E}_1(x^{T_{n-1}} y^{M_n} \mid T_{n-2}, M_{n-1})) \\
&= \sum_{i,j=0}^{\infty} \mathbb{P}_1(T_{n-2} = i, M_{n-1} = j) \sum_{k=j}^{\infty} \sum_{l=0}^{\infty} x^k y^l \mathbb{P}_1(T_{n-1} = k, M_n = l \mid T_{n-2} = i, M_{n-1} = j) \\
&= \sum_{i,j=0}^{\infty} \mathbb{P}_1(T_{n-2} = i, M_{n-1} = j) \sum_{k=j}^{\infty} \sum_{l=0}^{\infty} x^k y^l \mathbb{P}_j(T_0 = k, M_1 = l) \\
&= \sum_{i,j=0}^{\infty} \mathbb{P}_1(T_{n-2} = i, M_{n-1} = j)(\varphi(x, y))^j \\
&= \varphi_{n-1}(1, \varphi(x, y)),
\end{aligned}$$

where we use that $\{(T_{n-1}, M_n) : n \in \mathbb{N}\}$ is Markovian, Lemma 1.1 and the branching property.

3 The process conditioned to non-extinction

This section is devoted to study the process $\{(T_n, M_{n+1}), n \in \mathbb{Z}_+\}$ conditioned to non-extinction.

3.1 Construction

Here our aim is to prove Theorem 1.5, that ensures the existence of a Markovian process that we shall understand as the chain $\{(T_n, M_{n+1}), n \in \mathbb{Z}_+\}$, conditioned to the non-extinction of mutants in the population.

Proof of Theorem 1.5. An application of the monotone convergence Theorem along with an elementary computation shows that

$$\frac{d}{ds} \mathbb{E}_a(s^{M_n})\big|_{s=q} = \mathbb{E}_a\left(M_n q^{M_n-1}\right).$$

Moreover, the following identity is deduced form the branching property of the Galton-Watson process $\{M_n : n \in \mathbb{Z}_+\}$ and the properties of its generating function

$$\frac{d}{ds} \mathbb{E}_a(s^{M_n})\big|_{s=q} = aq^{a-1} f_n'(q).$$

The latter and former identities imply in turn that

$$\mathbb{E}_a\left(M_n q^{M_n-1}\right) = aq^{a-1} f_n'(q).$$

Then by the Markov property,

$$\mathbb{E}_a\left(M_{n+k} q^{M_{n+k}-1} \big| \mathcal{F}_n\right) = M_n q^{M_n-1} f_k'(q).$$

Combining the latter with the fact that $f_k'(q) = [f'(q)]^k$ (see Athreya and Ney [1], Lemma 3.3), we have that

$$Y_n = \frac{M_n q^{M_n-a}}{(f'(q))^n},$$

is martingale. Now from the theory of h-transform there exists a Markovian process (see [6]), that we denote by $\{(T_n^{\uparrow}, M_{n+1}^{\uparrow}), n \in \mathbb{Z}_+\}$ whose law satisfies

$$\mathbb{P}_a^{\uparrow}(T_0^{\uparrow} = i_0, M_1^{\uparrow} = j_1, \cdots, T_{n-1}^{\uparrow} = i_{n-1}, M_n^{\uparrow} = j_n) := \mathbb{P}_a(A_n) \frac{j_n q^{j_n-a}}{a(f'(q))^n}, \qquad (3.1)$$

where for every $n \in \mathbb{N}$, and

$$A_n = \{T_0 = i_0, M_1 = j_1, \cdots, T_{n-1} = i_{n-1}, M_n = j_n\}, \quad i_0, j_1, ..., i_{n-1}, j_n \in \mathbb{N}. \qquad (3.2)$$

∎

3.2 Conditional laws

This subsection is devoted to prove Theorem 1.6, in this aim the keystone will be the generating function of $\{(T_n, M_{n+1}) : n \in \mathbb{Z}\}$, hence some of the results given in section 2 will be neccesary.

Proof of Theorem 1.6.   i) We will first ensure that the generating function converges. For this end observe that $\{M_n > 0\}$ under the event $\{n < T < \infty\}$. Then for all $x, y \in [0, 1]$, we have the identity

$$\widehat{\varphi}_n(x, y) := \mathbb{E}_1(x^{T_{n-1}} y^{M_n} \mid n < T < \infty)$$
$$= \frac{\varphi_n(x, y) - \varphi_n(x, 0)}{1 - \mathbb{P}(M_n = 0)}.$$

Using the identity (2.6) and the fact $f_n(0) = \mathbb{P}(M_n = 0)$, we get the expression.

$$\widehat{\varphi}_n(x, y) = \frac{1 - f_{n-1}(0)}{1 - f_n(0)} \left(\frac{f_{n-1}(\varphi(x, y)) - f_{n-1}(0)}{1 - f_{n-1}(0)} - \frac{f_{n-1}(\varphi(x, 0)) - f_{n-1}(0)}{1 - f_{n-1}(0)} \right).$$

Now we take $u = f_{n-1}(0)$ and use that $m = f'(1)$ to obtain the following limit.

$$\lim_{n \to \infty} \frac{1 - f_{n-1}(0)}{1 - f_n(0)} = \lim_{u \to 1} \frac{1 - u}{1 - f(u)} = \frac{1}{m}.$$

Besides, note that for each s the function

$$n \longmapsto \frac{1 - f_{n-1}(s)}{1 - f_{n-1}(0)},$$

is a decreasing function. This implies that as n tends to infinity the expression

$$\frac{f_{n-1}(s) - f_{n-1}(0)}{1 - f_{n-1}(0)} = 1 - \frac{1 - f_{n-1}(s)}{1 - f_{n-1}(0)},$$

has a limit, say $1 - \widehat{f}(s)$. We stress (see for instance Athreya and Ney [1], Theorem 8.1) that $\widehat{f}(s)$ is the generating function of $\widehat{m} = \{\widehat{m}_k : k \in \mathbb{N}\}$, the Yaglom distribution of $\{M_n : n \in \mathbb{Z}_+\}$, which is obtained for all $k \in \mathbb{N}$, as follows

$$\widehat{m}_k = \lim_{n \to \infty} \mathbb{P}(M_n = k \mid n < T < \infty).$$

This fact implies that, the generating function $\widehat{f}$ satisfies

$$1 - \widehat{f}(f(s)) = m(1 - \widehat{f}(s)), \quad s \in [0, 1]. \tag{3.3}$$

The above calculations imply the following

$$\widehat{\varphi}(x, y) := \lim_{n \to \infty} \widehat{\varphi}_n(x, y) = \frac{\widehat{f}(\varphi(x, y)) - \widehat{f}(\varphi(x, 0))}{m}.$$

Now we prove by induction (1.7). If $n = 1$, it is the just proved equality. Then suppose (1.7) holds for $n = k$. In order to get the identity for $n = k + 1$ note that by the induction hypothesis

$$m^{k+1} \widehat{\varphi}(x, y) = m \left[1 - \widehat{f}(\varphi_k(x, 0)) \right] - m \left[1 - \widehat{f}(\varphi_k(x, y)) \right], \quad x, y \in [0, 1].$$

From where we deduce the claim by using first (3.3) and then (2.6).

9

ii) Let us consider the event of the form in (3.2). Since $\{(T_n, M_{n+1}), n \in \mathbb{Z}_+\}$ is Markovian we have for A_n as in (3.2), that

$$\mathbb{P}_a(A_n, n + k < T < \infty) = \mathbb{E}_a(\mathbf{1}_{A_n} \mathbb{E}_{j_n}(\mathbf{1}_{\{M_k>0\}} q^{M_k})),$$

and similarly

$$\mathbb{P}_a(n + k < T < \infty) = \mathbb{E}_a(\mathbf{1}_{\{M_{n+k}>0\}} q^{M_{n+k}}).$$

Then using (2.2), we get

$$\mathbb{P}_a(A_n \mid n + k < T < \infty) = \mathbb{P}_a(A_n) \frac{\sum_{j=1}^{\infty} \mathbf{P}^k_{(j_n,j)} q^j}{\sum_{j=1}^{\infty} \mathbf{P}^{n+k}_{(1,j)} q^j}.$$

Theorem 7.4 in [1] establishes that the following limit holds

$$\lim_{k \to \infty} \frac{\mathbf{P}^{n+k}_{(i_1,j)}}{\mathbf{P}^k_{(i_2,j)}} = i_1 i_2^{-1} (f'(q))^k q^{i_1 - i_2}.$$

Finally, thanks to hypothesis (H2) we can apply the previous result to obtain

$$\lim_{k \to \infty} \mathbb{P}_a(A_n | n + k < T < \infty) = \mathbb{P}_a(A_n) \frac{j_n q^{j_n - a}}{a(f'(q))^n}, \quad a \in \mathbb{N},$$

which finishes the proof.

∎

Remark 3.1. In the previous proof we established the existence of a Yaglom limit when $m \leq 1$, however similar arguments can be used to show that existences of a Yaglom limit in the supercritical case.

3.3 Interpretation

Motivated by the interpretation of a Galton-Watson process conditioned to non-extinction given in [11], in the present subsection we provide a description of $\{(T_n^{\uparrow}, M_{n+1}^{\uparrow}) : n \in \mathbb{Z}\}$ as immigration of mutants of the n-th type into the population. We start calculating the generating function of the n-step transition probabilities of the process.

Proposition 3.2. *The generating function of the n-step transition probabilities for the process*

$$\{(T_n^{\uparrow}, M_{n+1}^{\uparrow}), n \in \mathbb{Z}_+\}$$

is given by

$$\sum_{k,l=1}^{\infty} Q^n_{(i,j),(k,l)} x^k y^l = \frac{yq^{1-j}}{[f'(q)]^n} [\varphi_n(x, qy)]^{j-1} \frac{\partial}{\partial y} \varphi(x, qy) \prod_{i=1}^{n-1} f'(\varphi_i(x, qy)), \quad x, y \in [0, 1]. \quad (3.4)$$

Proof. Since $x, y \leq 1$ the generating function of (1.6) is infinitely differentiable, then an elementary calculation and the use of (2.5) imply the following identity.

$$\sum_{k,l=1}^{\infty} Q^n_{(i,j),(k,l)} x^k y^l = \frac{yq^{1-j}}{[f'(q)]^n} \left[\varphi_n^{j-1}(x,u) \frac{\partial}{\partial u} \varphi_n(x,u) \right]_{u=qy}.$$

To finish the proof we apply repeatedly (2.6) and the recursion $f_n(y) = f(f_{n-1}(y))$. ∎

Taking $x = 1$ in (3.4) and recalling the fact that the transition probabilities of $\{(T_n, M_{n+1}), n \in \mathbb{Z}_+\}$ depend only on the second coordinate, we can identify a Galton-Watson process with immigration [10].

Corollary 3.3. *If $\{M_n, n \in \mathbb{Z}_+\}$ is critical or subcritical, then $\{M_n^\uparrow - 1, n \in \mathbb{Z}_+\}$ is a Galton-Watson process with immigration $[f, \frac{f'}{m}]$.*

Note that $\{M_n^\uparrow, n \in \mathbb{Z}_+\}$ is the Q-process associated to the Galton-Watson process $\{M_n, n \in \mathbb{Z}_+\}$ (see for instance [11] or [1]). The following Corollary is analogous to Proposition 1 in [3].

Corollary 3.4. *If $\{M_n, n \in \mathbb{Z}_+\}$ is critical or subcritical, then the generating function of $(T_0^\uparrow, M_1^\uparrow)$ is given by the equation*

$$\mathbb{E}_1(x^{T_0^\uparrow} y^{M_1^\uparrow}) = \frac{xy}{m} \frac{\partial}{\partial y} g(\varphi(x,y), y), \quad x, y \in [0,1].$$

Moreover, the distribution of $(T_0^\uparrow, M_1^\uparrow)$ is given by

$$\mathbb{P}_a^\uparrow(T_0^\uparrow = k, M_1^\uparrow = l) = \frac{l}{mk} \pi_{k-a,l}^{*k}, \quad k \geq a \geq 1 \text{ y } l \geq 0.$$

Proof. Taking $n = 1$ in the equality (3.4),

$$\mathbb{E}_1(x^{T_0^\uparrow} y^{M_1^\uparrow}) = \frac{y}{m} \frac{\partial}{\partial y} \varphi(x,y).$$

Then the first identity is obtained with the substitution of (2.3). To get the second one, recall the definition of $\mathbb{P}^\uparrow$ given in (3.1), then use (2.4). ∎

We can now give an interpretation to the process $\{(T_n^\uparrow, M_{n+1}^\uparrow), n \in \mathbb{Z}_+\}$, in terms of a tree of alleles with immigration $\mathcal{A}^\uparrow = (\mathcal{A}_u^\uparrow : u \in \mathbb{U})$, this will provide a description of the genealogical structure in a population conditioned to non extinction. The key elements are the tree of alleles $\mathcal{A} = (\mathcal{A}_u : u \in \mathbb{U})$ and the above corollaries.

We start defining $\mathcal{A}_\varnothing^\uparrow = T_0^\uparrow$ that is, the total number of individuals without mutations into the population, then according to a distribution with generating function f'/m, a random number of individuals of the same genetic type arrive. We enumerate the $M_1^\uparrow$ allelic sub-populations of the first type beget by $T_0^\uparrow$ in decreasing order, with the convention that in the case of ties, sub-populations of the same size are ranked uniformly at random. Then, using Corollary 3.3 we choose uniformly at random one of the first type sub-families in the tree of alleles, removing it and replace it by a population of size $T_0^\uparrow$ which begets allelic subpopulation according to $M_1^\uparrow$, where $(T_0^\uparrow, M_1^\uparrow)$ is given by Corollary 3.4. We continue with the construction by iteration, $\mathcal{A}_{uj}^\uparrow$ is the size of the j-th sub-population allelic of type $|u| + 1$ which descend from the allelic sub-family indexed by the vertex u, then we choose one of the sub-families of the type $|u| + 1$ to replace it for one of size $T_0^\uparrow$ which begets allelic subpopulation according to $M_1^\uparrow$.

4 Asymptotic behavior

4.1 The α-stable case

Consider a Galton-Watson process $\{Z_n^+, n \in \mathbb{Z}_+\}$ with reproduction law ξ^+ denoted by π^+, that is

$$\pi_k^+ = \mathbb{P}(\xi^{(c)} + \xi^{(m)} = k), \quad k \in \mathbb{Z}_+.$$

Suppose that π^+ is critical and (1.9) holds. The mutations appear in the population according to (1.12) and we focus on the situation (1.14) described in the Introduction. Our goal in this section is to prove Theorem 1.8, for that end we start by describing the normalizing constant $r(n)$ appearing therein.

Lemma 4.1. *If condition (1.9) holds, then there exists $r(n) \in RV_\infty^{1/\alpha}$ such that*

$$r(n)\pi^+(ndy) \xrightarrow[n\to\infty]{} c_\alpha \frac{dy}{y^{1+\alpha}},$$

in the sense of vague convergence on $(0, \infty)$, where $c_\alpha = 1/\Gamma(3 - \alpha)$. In particular

$$\exp\left\{-t \int_{[0,\infty)} (1 - e^{-\lambda y} - \lambda y) r(n)\pi^+(ndy)\right\} \xrightarrow[n\to\infty]{} e^{-t\lambda^\alpha}.$$

The proof is an elementary application of standard results from the theory of Regular Variation (see e.g. [4] for background), but we include a proof in Appendix A for sake of completeness.

4.1.1 An approximation for the reproduction law

In order to link the asymptotic behaviour of the reproduction law of a typical individual with that of the joint distribution of clones and mutants, we first link their Laplace transform. Although in the present setting we use some ideas of the standard Tauberian-Abelian Theorem, we remark that it is not straightforward application of this theorem because in the present setting we consider sequences of measures indexed by the positive integers that change, unlike to the standard case where only the normalizing constants change.

Lemma 4.2. *For every positive integer n, let ϕ_n be the Laplace transform of $\xi^{(n)} = (\xi^{(cn)}, \xi^{(mn)})$ under the measure $\mathbb{P}^{p(n)}$. Assume $\{\lambda(n) : n \geq 0\}$ is a positive sequence such that $\lambda(n) \to 0$, as $n \to \infty$. Then*

$$\phi_n(\lambda(n), \theta) \sim \phi^+\left((1 - p(n))(1 - e^{-\lambda(n)}) + p(n)(1 - e^{-\theta})\right) \quad \text{with } n \to \infty, \forall \theta \geq 0, \qquad (4.1)$$

where ϕ^+ is the Laplace transform of $\xi^{(+)}$. In particular, ϕ_n^m, respectively ϕ_n^c, the Laplace transform of $\xi^{(mn)}$, respectively $\xi^{(cn)}$, satisfy

$$\phi_n^m(\theta) \sim \phi^+(p(n)(1 - e^{-\theta})),$$
$$[\phi_n^c(\lambda(n)) \sim \phi^+((1 - p(n))(1 - e^{-\lambda(n)}))], \quad \text{as } n \to \infty, \forall \theta \geq 0.$$

Proof. According to (1.12), conditionally to $\xi^{(+)} = k$ the distribution of $\xi^{(m)}$ is Binomial with parameter (k, p). This fact implies the following equality in law

$$(\xi^{(c)}, \xi^{(m)}) \overset{\mathcal{L}}{=} \sum_{i=1}^{\xi^{(+)}} (\mathbf{1}_{\{U_i > p\}}, \mathbf{1}_{\{U_i \leq p\}}), \tag{4.2}$$

where $\{U_i, i \geq 1\}$ are i.i.d. random variables with common distribution that of an uniform random variable in $(0, 1)$. Therefore,

$$\phi_n(\lambda(n), \theta) = \sum_{k=0}^{\infty} \mathbb{P}(\xi^{(+)} = k) \left[(1 - p(n))e^{-\lambda(n)} + p(n)e^{-\theta} \right]^k$$
$$= \phi^+(-\log(1 - (1 - (1 - p(n))e^{-\lambda(n)} - p(n)e^{-\theta}))).$$

We conclude the proof using (1.13) and the elementary asymptotic estimate

$$\frac{\log(1 - y)}{y} \xrightarrow[y \to 0]{} -1. \tag{4.3}$$

∎

In the same way it is posible to establish the following estimate.

Corollary 4.3. *For every positive integer n, let ψ_n be the characteristic function of $\xi^{(n)} = (\xi^{(cn)}, \xi^{(mn)})$ under the measure $\mathbb{P}^{p(n)}$. Then*

$$\psi_n(\lambda(n), \theta) \sim \phi^+((1 - p(n))(1 - e^{i\lambda(n)}) + p(n)(1 - e^{i\theta})), \quad as \ \lambda(n) \xrightarrow[n \to \infty]{} 0, \forall \theta \geq 0. \tag{4.4}$$

In particular ψ_n^m, respectively ψ_n^c, the characteristic function of $\xi^{(mn)}$ respectively of $\xi^{(cn)}$, satisfies

$$\psi_n^m(\theta) \sim \phi^+(p(n)(1 - e^{i\theta})),$$
$$[\psi_n^c(\lambda(n)) \sim \phi^+((1 - p(n))(1 - e^{i\lambda(n)}))],$$

as $n \to \infty$, $\lambda(n) \to 0$ and for all $\theta \geq 0$.

Proof. Similarly to previous Lemma, using (4.2) we have

$$\psi_n(\lambda(n), \theta) = \sum_{k=0}^{\infty} \mathbb{P}(\xi^{(+)} = k) \left[(1 - p(n))e^{i\lambda(n)} + p(n)e^{i\theta} \right]^k$$
$$= \sum_{k=0}^{\infty} \mathbb{P}(\xi^{(+)} = k) \exp\{k \log(1 - ((1 - p(n))(1 - e^{i\lambda(n)}) - p(n)(1 - e^{i\theta})))\}$$

To conclude we apply the asymptotic estimate (4.3). ∎

We can now use the results above to give an estimation for the reproduction measure.

Proposition 4.4. *For every positive integer n, let $\pi^{(cn)}$, $\pi^{(mn)}$ be the reproduction laws of $\xi^{(cn)}$ and $\xi^{(mn)}$, respectively. Assume $\{y(n) : n \geq 0\}$ is any sequence such that $y(n) \to \infty$ as $n \to \infty$. In the regime (1.9), the asymptotic behaviour of the tail distribution of $\xi^{(\cdot)}$ is given by*

$$\bar{\pi}^{(\cdot)}(y(n)) \sim c_\alpha \bar{\pi}^+ \left(y(n)/\mathbb{E}(\xi^{(\cdot)}) \right), \quad as\ n \to \infty. \tag{4.5}$$

where $(\cdot) = cn, mn$ and $c_\alpha = 1/\Gamma(3 - \alpha)$.

Proof. We prove the statement for clones, the mutants case is fully similar. First note that in the same way as in the proof of Lemma 4.1 (see the Appendix A),

$$\mu^{cn}(x) = \int_0^x s \bar{\pi}^{cn}(s) ds, \quad x \geq 0, \tag{4.6}$$

is a measure on $[0, \infty)$ with Laplace transform $\mathcal{L}_{\mu^{cn}}$ such that

$$\lambda^2 \mathcal{L}_{\mu^{cn}}(\lambda) = \mathbb{E}\left(1 - e^{-\lambda \xi^{(cn)}} - \lambda \xi^{(cn)} e^{-\lambda \xi^{(cn)}} \right), \quad \lambda \geq 0.$$

We now replace λ by a sequence $\{\lambda(n) : n \geq 0\}$ such that $\lambda(n) \to 0$ as $n \to \infty$,

$$\lambda(n)^2 \mathcal{L}_{\mu^{cn}}(\lambda(n)) = 1 - \phi_n^c(\lambda(n)) + \lambda(n)(\phi_n^c)'(\lambda(n)).$$

An estimate of the term ϕ_n^c follows from Lemma 4.2

$$\phi_n^c(\lambda(n)) \sim \phi^+((1 - p(n))(1 - e^{-\lambda(n)})), \quad n \to \infty.$$

In order to estimate $(\phi_n^c)'$ we use the fact that for every fixed n, conditionally to $\xi^{(+)} = k$ the distribution of $\xi^{(cn)}$ is Binomial with parameter $(k, 1 - p(n))$, and we proceed as in the Lemma 4.2 to get

$$(\phi_n^c)'(\lambda(n)) \sim (1 - p(n))e^{-\lambda(n)}\phi^{+'}((1 - p(n))(1 - e^{-\lambda(n)})), \quad n \to \infty.$$

Putting both terms together we infer the estimate,

$$\lambda(n)^2 \mathcal{L}_{\mu^{cn}}(\lambda(n)) \sim 1 - \phi^+((1 - p(n))(1 - e^{-\lambda(n)})) + (1 - p(n))(1 - e^{-\lambda(n)})\phi_n^{+'}((1 - p(n))(1 - e^{-\lambda(n)}))$$
$$- (1 - p(n))(1 - e^{-\lambda(n)} - \lambda(n)e^{-\lambda(n)})\phi_n^{+'}((1 - p(n))(1 - e^{-\lambda(n)})), \quad as\ n \to \infty.$$

Besides,

$$(1 - p(n))(1 - e^{-\lambda(n)} - \lambda(n)e^{-\lambda(n)})\phi_n^{+'}((1 - p(n))(1 - e^{-\lambda(n)})) \xrightarrow[n \to \infty]{} 0,$$

and from (A.2) we have

$$\lambda(n)^2 \mathcal{L}_\mu(\lambda(n)) = 1 - \phi^+(\lambda(n)) + \phi^{+'}(\lambda(n)), \quad n \to \infty,$$

where $\mathcal{L}_\mu$ is the Laplace transform of the measure μ defined in (A.1). From these last two displays we obtain

$$\lambda(n)^2 \mathcal{L}_{\mu^{cn}}(\lambda(n)) \sim ((1 - p(n))(1 - e^{-\lambda(n)}))^2 \mathcal{L}_\mu((1 - p(n))(1 - e^{-\lambda(n)})) + O(\lambda(n)^2), \quad as\ n \to \infty.$$

Due to the estimate $\lambda(n) \sim 1 - e^{-\lambda(n)}$, in the limit as $n \to \infty$, the approximation of $\mathcal{L}_\mu$ given in (A.3) implies

$$c_\alpha \lambda(n)^2 \mathcal{L}_{\mu^{cn}}(\lambda(n)) \sim \bar{\pi}^+ \left(\frac{1}{\lambda(n)(1 - p(n))} \right) + O(\lambda(n)^2), \quad n \to \infty. \tag{4.7}$$

Hence is remains to prove

$$\lim_{n \to \infty} \frac{\bar{\pi}^{cn}(1/\lambda(n))}{(\lambda(n))^2 \mathcal{L}_{\mu^{cn}}(\lambda(n))} = c_\alpha. \tag{4.8}$$

In this aim, define for every $y \geq 0$ the following measure

$$m^{cn}_{\lambda(n)}(0, y] := m^{cn}_{\lambda(n)}(y) = \frac{\mu^{cn}(y/\lambda(n))}{\mathcal{L}_{\mu^{cn}}(\lambda(n))}.$$

Observe that

$$\int_{[0,\infty)} e^{-\theta s} d_y \left(\frac{\mu^{cn}(y/\lambda(n))}{\mathcal{L}_{\mu^{cn}}(\lambda(n))} \right) = \frac{1}{\mathcal{L}_{\mu^{cn}}(\lambda(n))} \int_{[0,\infty)} e^{-\theta \lambda(n) y} \mu^{cn}(dy) = \frac{\mathcal{L}_{\mu^{cn}}(\theta \lambda(n))}{\mathcal{L}_{\mu^{cn}}(\lambda(n))}, \quad \forall \theta > 0,$$

from the previous display and (4.7), we get

$$\mathcal{L}_{m^{cn}_{\lambda(n)}}(\theta) \xrightarrow[n \to \infty]{} \theta^{-(2-\alpha)}, \quad \forall \theta > 0.$$

Writing now $\theta^{-(2-\alpha)}$ in terms of the gamma function we get that

$$\theta^{-(2-\alpha)} = \frac{1}{\Gamma(2-\alpha)} \int_0^\infty s^{(2-\alpha)-1} e^{-\theta s} ds,$$

and since the convergence of the Laplace ransform implies the weak convergence (see [7]), we have

$$m^{cn}_{\lambda(n)}(y) \xrightarrow[n \to \infty]{} \frac{1}{\Gamma(2-\alpha)} \int_0^y s^{(2-\alpha)-1} ds = \frac{y^{2-\alpha}}{\Gamma(3-\alpha)}. \tag{4.9}$$

Besides, by the definition of the measure $\mu^{(cn)}$ we can obtain that for any $y < 1$, the following inequality

$$(1/\lambda(n))^2 y(1-y) \bar{\pi}^{cn}(y/\lambda(n)) \leq \mu^{cn}(1/\lambda(n)) - \mu^{cn}(y/\lambda(n)) \leq (1/\lambda(n))^2 (1-y) \bar{\pi}^{cn}(1/\lambda(n)),$$

holds. Now, due to (4.9), we obtain

$$c_\alpha \frac{1 - y^{2-\alpha}}{1 - y} \leq \liminf_{n \to \infty} \frac{\bar{\pi}^{cn}(1/\lambda(n))}{(\lambda(n))^2 \mathcal{L}_{\mu^{cn}}(\lambda(n))} \leq \limsup_{n \to \infty} \frac{\bar{\pi}^{cn}(1/\lambda(n))}{(\lambda(n))^2 \mathcal{L}_{\mu^{cn}}(\lambda(n))} \leq c_\alpha \frac{1 - y^{2-\alpha}}{y(1-y)} \quad \text{for all } y < 1.$$

To conclude we make $y \uparrow 1$. ∎

15

4.2 Proof of Theorem 1.8

Before proving Theorem 1.8, let us we describe the scheme of proof that we will follow. According with the construction of the alleles tree $\{\mathcal{A}_u : u \in \mathbb{U}\}$, given $u \in \mathbb{U}$ such that $|u| = k \geq 1$, for all $j \in \mathbb{N}$ vertices uj, represent the size of the j-th allelic sub-populations of type $k+1$ begeted by u, thus labels of vertices at level $k+1$ determine the variable T_{k+1}. Moreover for all $k \in \mathbb{N}$, the total number of vertices at level k correspond to M_k. Besides a tree-indexed CSBP is related to the CSBP in the following way: vertices $u \in \mathbb{U}$ at level $|u| = k \geq 1$, represent the sizes of the sub-populations at generation k in the $CSBP$ which descend from the same parent at generation $k-1$. So, a first step to prove the converge of alleles trees in Theorem 1.8 is to prove the Proposition below.

Proposition 4.5. *Assuming (1.9) and (1.14), we have*

$$\mathcal{L}\left(\left(\left(\frac{T_k}{r(n)}, \frac{M_{k+1}}{r(n)p(n)}\right) : k \in \mathbb{Z}_+\right), \mathbb{P}^{p(n)}_{a(n)}\right) \Longrightarrow ((Z_{k+1}^{1/\alpha}, Z_{k+1}^{1/\alpha}) : k \in \mathbb{Z}_+),$$

where $(Z_k^{1/\alpha} : k \in \mathbb{Z}_+)$ is a CSBP process with reproduction measure ν^α given in Theorem 1.8.

Since the processes $\{(T_n, M_{n+1}) : n \in \mathbb{Z}\}$ is Markovian with transition probabilities given in (1.1), to get the previous convergence we need the following Lemma.

Lemma 4.6. *For $\alpha \in (1, 2)$, let τ_x^α be an $1/\alpha$-stable subordinator with Laplace exponent*

$$-\log \mathbb{E}(e^{-\lambda \tau_x^\alpha}) = x \int_0^\infty (1 - e^{\lambda y}) \nu^\alpha(dy), \tag{4.10}$$

where ν^α is as defined in Theorem 1.8. Assuming (1.9), (1.13) and (1.14)

i) the following convergences holds:

$$\mathcal{L}\left(\left(\frac{T_0}{r(n)}, \frac{M_1}{r(n)p(n)}\right), \mathbb{P}^{p(n)}_{a(n)}\right) \Longrightarrow (\tau_x^\alpha, \tau_x^\alpha); \tag{4.11}$$

ii) under the measure $\mathbb{P}_1^{p(n)}$, the behaviour of the joint tail distribution of (T_0, M_1), is given by

$$\lim_{n \to \infty} r(n)p(n)\mathbb{P}_1^{p(n)}\left(\frac{T_0}{r(n)} > s, \frac{M_1}{r(n)p(n)} > t\right) = \bar{\nu}^\alpha(s \wedge t), \tag{4.12}$$

where $\bar{\nu}^\alpha$ denotes the tail function of the Lévy measure ν^α.

Observe that the convergence in (4.11) establishes Poposition 4.5 in the case $k = 0$. Because a discrete time CSBP can be constructed from subordinators, that is, given a subordinator $\{\tau_t^\alpha : t \geq 0\}$, the sequence $\{Z_k : k \in \mathbb{Z}_+\}$ defined by iteration as follows:

$$\sum_{i=1}^{k+1} Z_i = \tau_{x+Z_1+\cdots+Z_k}^\alpha, \quad k \in \mathbb{Z}_+,$$

with the convention of $Z_0 = x$ for some fix $x > 0$, is a discrete time CSBP, we refer to [12] for further details.

The tree-indexed CSBP $\mathcal{Z}$ can also be constructed from this subordinator, namely according to [3], $\mathcal{Z}_\varnothing = x$ and $\{\mathcal{Z}_j : j \in \mathbb{N}\}$ is given by the sequence of the jumps of τ^α on $(0, x]$ ranked in decreasing order. The sizes of the subfamilies beget by the subfamily of size $\mathcal{Z}_1$, denoted by $\{\mathcal{Z}_{1j} : j \in \mathbb{N}\}$, correspond to the ranked sequences of the jumps of τ^α on the interval $(x, x + \tau^\alpha_{\mathcal{Z}_1}]$ ranked in decreasing order, $\{\mathcal{Z}_{2j} : j \in \mathbb{N}\}$ to those on the interval $(x + \tau^\alpha_{\mathcal{Z}_1}, x + \tau^\alpha_{\mathcal{Z}_1 + \mathcal{Z}_2}]$. Hence, the convergence (4.11) will establish the convergence in Theorem 1.8 with $u = \varnothing$, once the subordinator τ^α_x, is written in terms of their jumps, or by Lévy-Itô decomposition, in terms of the atoms of the Poisson point process associated with it. This is the aim of Lemma below.

Lemma 4.7. *Let $b(n)$ be a sequence of integers such that $b(n) \sim br(n)p(n)$ for some $b > 0$.*

a) *For every $n \in \mathbb{N}$, let $\{\chi^{(n)}_j : 1 \leq j \leq b(n)\}$ be a sequence of independent identically distributed random variables with distribution $\left(\frac{T_0}{r(n)}, \frac{M_1}{r(n)p(n)}\right)$. Defining for every $n \in \mathbb{N}$, $\gamma^{(n)}_j := \delta_{\chi^{(n)}_j}$ and $\gamma_n = \sum_{j=1}^\infty \gamma^{(n)}_j$, the following convergence holds*

$$\gamma_n \Longrightarrow \gamma, \tag{4.13}$$

where γ is a Poisson point process with intesity $b\eta$, with η the image of the measure ν^α (given in Theorem 1.8) by the action of the map $x \mapsto (x, x)$.

b) *We have the following convergence, under the measure $\mathbb{P}_1^{p(n)}$*

$$\left(\frac{T_0}{r(n)}, \frac{M_1}{r(n)p(n)}\right)^{(b(n)^\downarrow)} \Longrightarrow (\mathbf{a}_1, \mathbf{a}_2, ...), \tag{4.14}$$

where for all $k \in \mathbb{N}$, $\mathbf{a}_k = (a_k, a_k)$ with $\{a_k\}_{k \in \mathbb{N}}$, the atoms of a Poisson random measure on $(0, \infty)$ with intensity $b\nu^\alpha$ ranked in decreasing order; the measure ν^α is given in Theorem 1.8.

Proof. From the construction $\{\gamma^{(n)}_j : 1 \leq j \leq b(n)\}$ is a sequence of independent random variables. Besides, the convergence (4.12) in Lemma 4.6 implies $\gamma^{(n)}_j \xrightarrow{P} 0$ as $n \to \infty$, uniformly in j. Then

$$\sup_j \mathbb{E}(|\gamma^{(n)}_j \wedge 1|) \to 0.$$

Hence, according with the definition given in chapter 4 of [9], we have that $\{\gamma^{(n)}_j : 1 \leq j \leq b(n)\}$ is a null array. Thus we will get the convergence (4.13) as an application of Theorem 16.18 of [9], once we verify the following condition:

i) $\sum_j \mathbb{P}(\gamma^{(n)}_j(B) > 0) \to \eta(B)$, as $n \to \infty$, for all $B \in \hat{\mathcal{B}}$, where η is the image of the measure $b\nu^\alpha$ by the application $x \mapsto (x, x)$,

ii) $\sum_j \mathbb{P}(\gamma^{(n)}_j(B) > 1) \to 0$, as $n \to \infty$, for all $B \in \mathcal{B}$.

where $\mathcal{B}$ is the Borel σ-algebra of $[0, \infty)^2$, $\hat{\mathcal{B}} =: \{B \in \mathcal{B}, \gamma(\partial B) = 0 \text{ c.s.}\}$, with γ the measure defined in the statement of this Lemma and the symbol ∂ denotes the boundary of B. Observe that the class of sets $B = ((b, \infty) \times \mathbb{R}_+) \cup (\mathbb{R}_+ \times (b', \infty))$ is a π-system which generates a λ-system

17

and it is just $\mathcal{B}$. Then by Dynkin's Theorem, is enough to establish the conditions above for sets of the form B. In this setting, the condition (ii) holds because $\gamma_j^{(n)}(\cdot)$ takes only the values 0 or 1, for any j and n, the condition (ii) is true. To establish (i), observe the following equalities

$$\sum_{j=1}^{b(n)} \mathbb{P}(\gamma_j^{(n)}(B) > 0) = \sum_{j=1}^{b(n)} \mathbb{P}(\gamma_j^{(n)}(B) = 1)$$

$$= \sum_{j=1}^{b(n)} \bar{\mu}_n((r(n))^{-1}s, (r(n)p(n))^{-1}t)$$

$$= b(n)\bar{\mu}_n((r(n))^{-1}s, (r(n)p(n))^{-1}t),$$

here we recall the identity $\bar{\mu}_n(x,y) = \mathbb{P}_1^{p(n)}(T_0 > x \text{ or } M_1 > y)$. Assuming that $b(n) \sim br(n)p(n)$ for some $b > 0$, from Lemma 4.6 and the last equality we have

$$\sum_{j=1}^{b(n)} \mathbb{P}(\gamma_j^{(n)}(B) > 0) \xrightarrow[n\to\infty]{} b\bar{\nu}^\alpha(s \wedge t)).$$

To get the first convergence in the Lemma, it remains to observe that the equality $b\bar{\nu}^\alpha(s \wedge t) = \eta(B)$ holds. But this follows from the following equalities,

$$\int_B \eta(dx, dy) = b \int_{(x,x)\in B} \nu^\alpha(dx) = b \int_{(s\wedge t,\infty)} \nu^\alpha(dx).$$

We will now prove the convergence (b). For $i = 1, 2$, $\chi_{ij}^{(n)}$ denotes the i-th coordinate of the sequence $\chi_j^{(n)}$ that appear in the statement (a). Assuming that $\chi_i^{(n)} := (\chi_{1i}^{(n)}, \chi_{2i}^{(n)}) \leq \chi_j^{(n)} := (\chi_{1j}^{(n)}, \chi_{2j}^{(n)})$ if and only if $\chi_{1i}^{(n)} \leq \chi_{1j}^{(n)}$ or $\chi_{2j}^{(n)} \leq \chi_{2j}^{(n)}$, let us define j_1 as the index where the maximum integer of the sequence $\{\chi_i^{(n)} : 1 \leq i \leq b(n)\}$ is reached.

$$\chi_{j_1}^{(n)} = \max_{1 \leq i \leq b(n)} \chi_i.$$

Similarly for $k = 2, ..., b(n)$, let j_k be the index of the k-order statistic

$$\chi_{j_k}^{(n)} = \max_{i \in J_k} \chi_i,$$

where $J_k = \{1, ..., b(n)\} \setminus \{j_1, ..., j_{k-1}\}$. Then observe that

$$\mathbb{P}(\chi_{j_1}^{(n)} \geq \mathbf{c}_1, \chi_{j_2}^{(n)} \geq \mathbf{c}_2, ..., \chi_{j_k}^{(n)} \geq \mathbf{c}_k) = \mathbb{P}(\gamma_n(C_1) \geq 1, \gamma_n(C_2) \geq 2, ..., \gamma_n(C_k) \geq k),$$

if $\mathbf{c}_i = (c_i, c_i)$, $C_i = (0,1) \times (0,1) \setminus (0,c_i) \times (0,c_i)$ and $c_1 > \cdots > c_k$. Taking now the limit when $n \to \infty$ in the equality below and using the convergence in (4.12), we have

$$\mathbb{P}(\chi_{j_1} \geq \mathbf{c}_1, \chi_{j_2} \geq \mathbf{c}_2, ..., \chi_{j_k} \geq \mathbf{c}_k) \xrightarrow[n\to\infty]{} \mathbb{P}(\gamma(C_1) \geq 1, \gamma(C_2) \geq 2, ..., \gamma(C_k) \geq k).$$

This implies the desired convergence because

$$\mathbb{P}(\gamma(C_1) \geq 1, ..., \gamma(C_k) \geq k) = \mathbb{P}(\mathbf{a}_{j_1} \geq \mathbf{c}_1, \mathbf{a}_{j_2} \geq \mathbf{c}_2, ..., \mathbf{a}_{j_k} \geq \mathbf{c}_k),$$

where $\mathbf{a}_k = (a_k, a_k)$ with $\{a_k\}_{k\in\mathbb{N}}$, the atoms of a Poisson random measure on $(0, \infty)$ with intensity $b\nu^\alpha$ ranked in decreasing order; the measure ν^α is given in Theorem 1.8. As before we used the indices j_k to rank in decreasing order the sequence $\mathbf{a}_k$. ∎

4.2.1 Proof of Lemma 4.6

For each $n \in \mathbb{N}$, let $\mathbf{S}_k^{(n)}$ be a random vector such that their coordinates are obtained from $\xi^{(cn)}$ and $\xi^{(mn)}$, as follows

$$\mathbf{S}_k^{(n)} = (a(n), 0) + \sum_{i=1}^{k} \left(\xi_i^{(cn)} - 1, \xi_i^{(mn)} \right), \quad k \in \mathbb{N},$$

where $\{ (\xi_i^{(cn)}, \xi_i^{(mn)}) : i \in \mathbb{N} \}$ are a pair of i.i.d. random variables with distribution $(\xi^{(cn)}, \xi^{(mn)})$.

A key result in the proof of Lemma 4.6 is the following:

Proposition 4.8. *In the regime (1.9), (1.13) and (1.14) the normalized random walk defined by*

$$\bar{\mathbf{S}}_{\lfloor r(n)t \rfloor}^{(n)} = (a(n)/r(n)p(n), 0) + \sum_{i=1}^{\lfloor r(n)t \rfloor} \left((\xi_i^{(cn)} - 1)/n, \xi_i^{(mn)}/r(n)p(n) \right), \quad t \geq 0,$$

converges weakly

$$\left(\bar{\mathbf{S}}_{\lfloor r(n)t \rfloor}^{(n)} : t \geq 0 \right) \Longrightarrow ((x + X_t, t) : t \geq 0),$$

where $\{X_t, t \geq 0\}$ is an α-stable process with no-negative jumps with and characteristic exponent $c_\alpha |\lambda|^\alpha$.

We pospone the proof of this result to the appendix and focus on the proof of Lemma 4.6.

Proof of Lemma 4.6 i) . From Lemma 3 of [3], we know that the first passage time below 0 for the centered random walk $\mathbf{S}_{1,k}^{(n)} = a(n) + \sum_{i=1}^{k} (\xi_k^{(cn)} - 1)$ has the same distribution as T_0. Moreover, on the one hand, because of the identities,

$$\varsigma^{(n)}(0) = \inf\{k \in \mathbb{Z}_+ : \mathbf{S}_{1,k}^{(n)} = 0\} \quad \text{and} \quad \Sigma^{(n)}(0) = \sum_{i=1}^{\varsigma^{(n)}(0)} \xi_i^{(mn)},$$

the random variables

$$(\varsigma^{(n)}(0), \Sigma^{(n)}(0)) \quad \text{and} \quad (T_0, M_1),$$

have the same distribution under $\mathbb{P}_{a(n)}^{p(n)}$. On the other hand, we also have the following two identities

$$\frac{\varsigma^{(n)}(0)}{r(n)} = \frac{1}{r(n)} \inf\{k \in \mathbb{Z}_+ : \mathbf{S}_{1,k}^{(n)} = 0\} = \inf\{t \geq 0 : \bar{\mathbf{S}}_{1, \lfloor r(n)t \rfloor}^{(n)} = 0\}.$$

and

$$\left(\frac{1}{r(n)} \varsigma^{(n)}(0), \bar{\mathbf{S}}_{\varsigma^{(n)}(0)}^{(n)} \right) = \left(\frac{1}{r(n)} \varsigma^{(n)}(0), \left(\bar{\mathbf{S}}_{1, \lfloor \varsigma^{(n)}(0) \rfloor}^{(n)}, \frac{1}{r(n)p(n)} \Sigma^{(n)}(0) \right) \right), \quad (4.15)$$

and the weak convergence

$$\left(\bar{\mathbf{S}}_{\lfloor r(n)t \rfloor}^{(n)} : t \geq 0 \right) \Longrightarrow ((x + X_t, t) : t \geq 0).$$

19

Since X is a spectrally positive Lévy process, Theorem 1 in Chapter VII of [2] ensures that the first passage time below $-x$ for the process X

$$\tau_x^\alpha = \inf\{t \geq 0 : X_t \leq -x\}, \quad x \geq 0.$$

is a stable subordinator of parameter $1/\alpha$. We conclude from these facts that the claimed convergence holds as soon as

$$\left(\frac{1}{r(n)}\varsigma^{(n)}(0), \bar{\mathbf{S}}_{\varsigma^{(n)}(0)}^{(n)}\right) \Longrightarrow \left(\tau_x^\alpha, (x + X_t, t)|_{t=\tau_x^\alpha}\right). \tag{4.16}$$

But according to Theorem 13.6.5 of [14] about weak convergence of first passage times and under shoots and overshoots, when there is convergence in Skorohod's topology, we get that

$$\left(\frac{1}{r(n)}\varsigma^{(n)}(0), \bar{\mathbf{S}}_{1,\varsigma^{(n)}(0)}^{(n)}\right) \Longrightarrow \left(\tau_x^\alpha, X_{\tau_x^\alpha} + x\right)$$

Moreover since we have the joint convergence

$$\left(\bar{\mathbf{S}}_{1,\lfloor\varsigma^{(n)}(0)t\rfloor}^{(n)}, \bar{\mathbf{S}}_{2,\lfloor\varsigma^{(n)}(0)t\rfloor}^{(n)} : t \geq 0\right) \Longrightarrow \left((x + X_t, t) : t \geq 0\right)$$

in the sense of Skorohod's topology, and the second coordinate is a determinist linear function, its is an elementary exercise to extend the above mentioned result of [14] to get that the convergence in (4.16) holds. ∎

Proof of Lemma 4.6 ii). We will apply the same techniques used in the proof of statement (ii) in Lemma 4 of [3]. Let us start observing that for every $x, y \in \mathbb{R}$

$$e^{-sx-ty} = st\int_0^\infty \int_0^\infty e^{-sx-ty}\mathbf{1}_{\{x<u,y<v\}}dudv, \quad s, t \geq 0.$$

Thus Fubini's Theorem implies that for any random vector (X, Y) the following identity holds.

$$1 - \mathbb{E}(e^{-sX-tY}) = st\int_0^\infty \int_0^\infty e^{-sx-ty}\mathbb{P}(X \geq u \text{ or } Y \geq v)dudv, \quad s, t \geq 0.$$

In particular,

$$1 - \mathbb{E}_1^{p(n)}\left(e^{-\frac{s}{r(n)}T_0 - \frac{t}{r(n)p(n)}M_1}\right) = st\int_0^\infty \int_0^\infty e^{-su-tv}\bar{\mu}_n((r(n))^{-1}u, (r(n)p(n))^{-1}v)dudv, \quad s, t \geq 0,$$

where $\bar{\mu}_n(x, y) := \mathbb{P}_1^{p(n)}(T_0 > x \text{ or } M_1 > y)$. Hence by the branching property,

$$\mathbb{E}_{a(n)}^{p(n)}\left(e^{-\frac{s}{r(n)}T_0 - \frac{t}{r(n)p(n)}M_1}\right) = \left(1 - st\int_0^\infty \int_0^\infty e^{-su-tv}\bar{\mu}_n((r(n))^{-1}u, (r(n)p(n))^{-1}v)dudv\right)^{a(n)}. \tag{4.17}$$

Now, according to (i) the left term in the previous display converges as $n \to \infty$ towards

$$\mathbb{E}\left(e^{-(s+t)\tau_x^\alpha}\right).$$

20

Besides the right term converges as $n \to \infty$ towards

$$\exp\left(-x \int_0^\infty (1 - e^{-(s+t)y}) \nu^\alpha(dy)\right),$$

here we use (i) in the Lemma 4.6, together with (1.13). Therefore

$$\mathbb{E}\left(e^{-(s+t)\tau_x^\alpha}\right) = \exp\left(-x \int_0^\infty (1 - e^{-(s+t)y}) \nu^\alpha(dy)\right), \tag{4.18}$$

Taking logarithms in (4.17) and (4.18) we obtain

$$\lim_{n\to\infty} sta(n) \int_0^\infty \int_0^\infty e^{-su-tv} \bar\mu_n((r(n))^{-1}u, (r(n)p(n))^{-1}v) du dv = x \int_0^\infty (1 - e^{-(s+t)y}) \nu^\alpha(dy).$$

Hence it only remains to see that the line above is equal to

$$xst \int_0^\infty \int_0^\infty e^{-sy-tz} \bar\nu^\alpha(y \wedge z) dy dz.$$

For that end we observe the equality

$$\int_0^\infty \int_0^\infty e^{-sy-tz} \bar\nu^\alpha(y \wedge z) dy dz = \int_0^\infty \nu^\alpha(du) \int_0^\infty \int_0^\infty e^{-sy-tz} \mathbf{1}_{\{u>y \text{ or } u>z\}} dy dz,$$

and we obtain the claimed identity by uniqueness of the Laplace transform. ∎

4.2.2 Proof of Proposition 4.5

Observe it is enough to show the convergence of the Laplace transforms associated with the finite dimensional distributions of each processes. So, we will briefly deduce the Laplace transform of a CSBP (in discrete time) $\{Z_k : k \in \mathbb{Z}_+\}$. From [12], we know that the transition probabilities of this branching process are characterized, for every $k \in \mathbb{Z}_+$ and $\lambda, z \geq 0$ by

$$\mathbb{E}(e^{-\lambda Z_{k+1}}|Z_k = z) = e^{-z\kappa(\lambda)}. \tag{4.19}$$

where κ is the cumulant of a subordinator without drift, that is

$$\kappa(\lambda) = \int_{(0,\infty)} (1 - e^{-\lambda x}) \vartheta(dx), \tag{4.20}$$

where ϑ is the Lévy measure. Applying successively the property (4.19), we obtain

$$\mathbb{E}_x(e^{-s_1 Z_1 - \cdots - s_k Z_k}) = e^{-x\kappa(l_{k-1}(s_1))} \qquad s_i \geq 0, \; i = 1, 2, ..., k.$$

where by definition $l_0(s) = s$ and

$$l_i(s_{n-i}) = s_{n-i} + \kappa(l_{i-1}(s_{n-i+1})), \quad i \in \mathbb{Z}_+. \tag{4.21}$$

21

In particular, $\{Z_k^{1/\alpha} : k \in \mathbb{Z}_+\}$ is a CSBP whose transition probabilities are characterized by a subordinator τ_x^α, with cumulant given by (4.20) with $\vartheta = \nu^\alpha$ defined in Theorem 1.8. Therefore we will prove by induction on k the following convergence:

$$\mathbb{E}_{a(n)}^{p(n)}\left(\prod_{i=1}^{k} e^{-\frac{s_{i-1}}{r(n)}T_{i-1}-\frac{t_i}{r(n)p(n)}M_i}\right) \xrightarrow[n\to\infty]{} \exp\{-x\kappa(l_{k-1}(s_0+t_1))\}, \quad \text{for all } s_i, t_i \geq 0, i = 1, 2, ..., k.$$
(4.22)

where l is defined in (4.21), taking $\vartheta = \nu^\alpha$. We start observing that $Z_1^{1/\alpha}$ has the same distribution of τ_x^α because of CSBP's construction. Thus the case $n = 1$ follows from Lemma 4.6 (i). We next assume (4.22) holds for $n = k$ and prove the convergence for $n = k + 1$. To this end, we use the Markov property of $\{(T_n, M_{n+1}) : n \in \mathbb{Z}\}$ and the fact that, conditionally to $M_n = j$, the pair (T_n, M_{n+1}) has the same distribution of (T_0, M_1) under $\mathbb{P}_j$, to obtain

$$\mathbb{E}_{a(n)}^{p(n)}\left(\prod_{i=1}^{k+1} e^{-\frac{s_{i-1}}{r(n)}T_{i-1}-\frac{t_i}{r(n)p(n)}M_i}\right)$$

$$= \mathbb{E}_{a(n)}^{p(n)}\left(e^{-\frac{s_0}{r(n)}T_0-\frac{t_1}{r(n)p(n)}M_1}\cdots e^{-\frac{s_{k-1}}{r(n)}T_{k-1}-\left(\frac{t_k}{r(n)p(n)}-\frac{1}{r(n)p(n)}\log \mathbb{E}_{r(n)p(n)}^{p(n)}\left(e^{-\frac{s_k}{r(n)}T_0-\frac{t_{k+1}}{r(n)p(n)}M_1}\right)\right)M_k}\right).$$

Due to the assumption $r(n)p(n) \sim a(n)x$ in hypothesis (1.14), we obtain as consequence of Lemma 4.6 (i), that

$$\mathbb{E}_{a(n)}^{p(n)}\left(\prod_{i=1}^{k+1} e^{-\frac{s_{i-1}}{r(n)}T_{i-1}-\frac{t_i}{r(n)p(n)}M_i}\right) \sim \mathbb{E}_{a(n)}^{p(n)}\left(e^{-\frac{s_0}{r(n)}T_0-\frac{t_1}{r(n)p(n)}M_1}\cdots e^{-\frac{s_{k-1}}{r(n)}T_{k-1}-\frac{1}{r(n)p(n)}(t_k+\kappa(s_k+t_{k+1}))M_k}\right).$$

Then using the induction hypothesis with

$$s_{i-1}' + t_i' = \begin{cases} s_{i-1} + t_i & i < k, \\ l(s_{i-1} + t_i) & i = k, \end{cases}$$
(4.23)

we get

$$\mathbb{E}_{a(n)}^{p(n)}\left(e^{-\frac{s_0}{r(n)}T_0-\frac{t_1}{r(n)p(n)}M_1}\cdots e^{-\frac{s_{k-1}}{r(n)}T_{k-1}-\frac{1}{r(n)p(n)}(t_k+\kappa(s_k+t_{k+1}))M_k}\right)$$

$$\xrightarrow[n\to\infty]{} \exp\{-x\kappa(l_{k-1}(s_0'+t_1'))\}.$$

This concludes the proof because of the recursive definition of l_i given in (4.21) together with the choice of $s_{i-1}' + t_i'$,

$$\kappa(l_{k-2}(s_1'+t_2')) = \kappa(l_{k-1}(s_1+t_2)),$$

as consequence of $l_{k-1}(s_0'+ct_1') = l_k(s_0+t_1)$.

∎

22

4.2.3 Proof of Theorem 1.8

We will establish

$$\mathcal{L}\left(\left(\left(\left(\frac{\mathcal{A}_u^{(n)}}{r(n)}, \frac{d_u^{(n)}}{r(n)p(n)}\right) : |u| \leq k\right) : k \in \mathbb{Z}_+\right), \mathbb{P}_{a(n)}^{p(n)}\right) \Longrightarrow \left(\left((\mathcal{Z}_u, \mathcal{Z}_u) : |u| \leq k\right) : k \in \mathbb{Z}_+\right). \tag{4.24}$$

Actually by the monotone class theorem, it is enough to show for non-negative measurable continuous functions $f_1, ..., f_k$

$$\mathbb{E}_{a(n)}^{p(n)}\left[\prod_{i=1}^k f_i((r(n))^{-1}\mathcal{A}_u^{(n)}, (r(n)p(n))^{-1}d_u^{(n)} : |u| \leq i)\right] \xrightarrow[n\to\infty]{} \mathbb{E}^{\mathbb{Q}_x}\left[\prod_{i=1}^k f_i(\mathcal{Z}_u, \mathcal{Z}_u : |u| \leq i)\right],$$

where $\mathbb{Q}_x$ is the law of a tree-indexed CSBP started with an initial population of size x contructed from a subordinator $\{\tau_t^\alpha : t \geq 0\}$. Let us prove it by induction on k. The case $k = 1$ is given in the convergence (4.14). Assuming the case k, we will prove the convergence for $k + 1$. Conditioning with respect to $\mathcal{F}_k = \sigma(\mathcal{A}_u^{(n)}, d_u^{(n)} : |u| \leq k)$ and using Lemma 1.4 we have

$$\mathbb{E}_{a(n)}^{p(n)}\left[\mathbb{E}_{a(n)}^{p(n)}\left(\prod_{i=1}^{k+1} f_i((r(n))^{-1}\mathcal{A}_u^{(n)}, (r(n)p(n))^{-1}d_u^{(n)} : |u| = i) \mid \mathcal{F}_k\right)\right]$$

$$= \mathbb{E}_{a(n)}^{p(n)}\left[\prod_{i=1}^k f_i((r(n))^{-1}\mathcal{A}_u^{(n)}, (r(n)p(n))^{-1}d_u^{(n)} : |u| \leq i)\right.$$

$$\left. \mathbb{E}_1^{p(n)}\left(f_{k+1}(((r(n))^{-1}T_0^{(n)}, (r(n)p(n))^{-1}M_1^{(n)})^{d_u^{(n)}\downarrow} : |u| = k)\right)\right]$$

Besides, by induction hypothesis $d_u^{(n)} \sim r(n)p(n)\mathcal{Z}_u$ with $|u| = k$, therefore when $n \to \infty$ in the previous equality we obtain

$$\mathbb{E}^{\mathbb{Q}_x}\left[\prod_{i=1}^k f_i(\mathcal{Z}_u, \mathcal{Z}_u : |u| \leq i)\mathbb{E}_1(f_{k+1}((\mathbf{a}_1', \mathbf{a}_2', ...)))\right],$$

where $\mathbf{a}_k' = (a_k', a_k')$ are the atoms of a Poisson random measure on $(0, \infty)$ with intensity $b\mathcal{Z}_u\nu^\alpha$, repeated according to their multiplicity and ranked in the decreasing order. Due to the definition of a CSBP tree this concludes the proof.

∎

4.3 The conditioned to non-extinction case

This section is devoted to prove Theorem 1.9. Following the same strategy of Proposition 4.5, we shall establish by induction the convergence of Laplace transforms of the finite dimensional distributions associated with the processes involved. In this aim, we firstly deduce the Laplace transform of the finite dimensional distributions of a CSBP with immigration, with mechanism (ϑ, ι), $\{Z_n^I : n \in \mathbb{N}\}$, defined for every $n \in \mathbb{N}$ as follows

$$Z_{n+1}^I = \tau_n(Z_n^I) + I_n,$$

where $\{I_n : n \in \mathbb{N}\}$ is a sequence of nonnegative random variables with common probability measure ω, which determine the distribution of individual immigrants arriving in the population. Let us denote its Laplace transform of ω by ι, i.e.

$$\iota(\lambda) = \int_0^\infty e^{-\lambda x} \omega(dx), \quad \lambda \geq 0. \tag{4.25}$$

Let $\{T_n(t) : t \geq 0\}_{n \geq 0}$ be a sequence of independent subordinators (without drift) and also independents of I_n, with the same distribution and Laplace transform given in (4.20). Thereby

$$\mathbb{E}(e^{-\lambda T_n(Z_n^I)} | Z_n^I) = e^{-Z_n^I \kappa(\lambda)}.$$

This previous equality together with the Markov property imply that

$$\mathbb{E}_x(e^{-s_1 Z_1^I - \cdots - s_k Z_k^I}) = \prod_{i=0}^{k-1} \iota(l_i(s_{k-i})) e^{-x \kappa(l_{k-1}(s_1))}, \quad \text{for all } s_i \geq 0, i = 1, 2, ..., k.$$

Thus the proof of statement (i) of Theorem 1.9 require to establish for all $s_i, t_i \geq 0$, $i = 1, 2, ..., k$, the convergence below holds

$$\mathbb{E}_{a_n}^{p(n)\uparrow} \left(\prod_{i=1}^{k} e^{-\frac{s_{i-1}}{n^2} T_{i-1} - \frac{t_i}{n} M_i} \right) \xrightarrow[n \to \infty]{} \exp\{-x\kappa(l_{k-1}(s_0 + c_\vartheta t_1))\} \prod_{i=0}^{k-1} \iota(l_{k-i}(s_{i-1} + \tilde{c}_\vartheta t_i)), \tag{4.26}$$

where κ and l are respectively defined in (4.20) and (4.21), taking in partiular $\vartheta = \nu$ given in (1.16). To obtain (ii) of Theorem 1.9, the previous convergence is proved with $\vartheta = \nu^\alpha$ defined in (1.15). The following Lemma establishes the above convergence in the case $k = 1$. In their proof we use the reference [3] to justify the statement corresponding to $\vartheta = \nu$ and previous results for the associated to $\vartheta = \nu^\alpha$.

Lemma 4.9. *If (1.13) holds, then we have the following convergence*

$$\mathcal{L}\left((b_1(n)T_0, b_2(n)M_1), \mathbb{P}_{a(n)}^{p(n)\uparrow}\right) \implies (\tau, c_\vartheta \tau),$$

where τ is a random variable with Laplace transform $e^{-\kappa(s)} \iota(s)$, where $\kappa(s)$, $\iota(s)$ are given in (4.20) and (4.25), according to

a) *if the reproduction law has finite variance σ^2, $\vartheta = c^{-1}\nu$, where ν is the measure in (1.16). Moreover $b_1(n) = n^{-2}$, $b_2(n) = n^{-1}$ and $c_\vartheta = c$;*

b) *otherwise, under the assumptions (1.9) and (1.14), $b_1(n) = (r(n))^{-1}$, $b_2(n) = (r(n)p(n))^{-1}$, ϑ is given by (1.15) and $c_\vartheta = 1$.*

Proof. To simplify the notation we just write b_1 and b_2. We prove the convergence of Laplace transform $(b_1 T_0, b_2 M_1)$ under the measure $\mathbb{P}_{a(n)}^{p(n)\uparrow}$. First, recalling the definition of the conditional measure given in (3.1), an elementary calculation using the branching property implies

$$\mathbb{E}_a^{p\uparrow}\left(e^{-sT_0 - tM_1}\right) = \mathbb{E}_{a-1}^p\left(e^{-sT_0 - tM_1}\right) \mathbb{E}_1^p\left(e^{-sT_0 - tM_1} M_1\right). \tag{4.27}$$

Thanks to Lemma 4 of [3] and Lemma 4.6.

$$\mathbb{E}_{a(n)}^{p(n)}\left(\exp\left(-sb_1 T_0 - tb_2 M_1\right)\right) \xrightarrow[n\to\infty]{} \exp\left(-x\int_0^\infty (1 - e^{-(s+c_\vartheta t)y})c_\vartheta^{-1}\vartheta(dy)\right), \tag{4.28}$$

so it remains to calculate the limit of the second factor in (4.27). Using again [3] together with the equality (4.17), we have

$$\mathbb{E}_1^{p(n)}\left(\exp\left(-sb_1 T_0 - tb_2 M_1\right)\right) = 1 - st\int_0^\infty \int_0^\infty e^{-sx-ty}\bar{\mu}_n(b_1^{-1}x, b_2^{-1}y)dxdy,$$

as before $\bar{\mu}_n(x,y) := \mathbb{P}_1^{p(n)}(T_0 > x \text{ or } M_1 > y)$. This implies due to Lemma 4 (ii) in [3] and Lemma 4.6 (ii)

$$\mathbb{E}_1^{p(n)}\left(\exp\left(-sb_1 T_0 - tb_2 M_1\right) M_1\right) \xrightarrow[n\to\infty]{} s\int_0^\infty \int_0^\infty e^{-sx-ty}\bar{\vartheta}\left(x \wedge \frac{y}{c_\vartheta}\right)dxdy$$
$$- st\int_0^\infty \int_0^\infty y e^{-sx-ty}\bar{\vartheta}\left(x \wedge \frac{y}{c_\vartheta}\right)dxdy.$$

Computing the integrals we get

$$\mathbb{E}_1^{p(n)}\left(\exp\left(-sb_1 T_0 - tb_2 M_1\right) M_1\right) \xrightarrow[n\to\infty]{} \int_0^\infty e^{-(s+c_\vartheta t)z}z\vartheta(dz). \tag{4.29}$$

This finishes the proof. $\blacksquare$

We can now continue with the proof of Theorem 1.9, assuming that (4.26) holds for k and to verify the case $k+1$. Let be $\mathcal{F}_k = \sigma((M_{j-1}, T_j), j \leq k)$ and (T_0', M_1') an independent copy of (T_0, M_1). Recalling the definition of the measure $\mathbb{P}^\uparrow$ we have,

$$\mathbb{E}_a^{p\uparrow}\left(e^{-\lambda_0 T_0 - \theta_1 M_1}\cdots e^{-\lambda_k T_k - \theta_{k+1}M_{k+1}}\right)$$
$$= \mathbb{E}_a^p\left(e^{-\lambda_0 T_0 - \theta_1 M_1}\cdots e^{-\lambda_{k-1}T_{k-1} - \theta_k M_k}\frac{1}{a}\mathbb{E}_{M_k}^p\left(e^{-\lambda_k T_0' - \theta_{k+1}M_1'}M_1'\right)\right).$$

Then applying the identity (4.27) we get

$$\mathbb{E}_a^{p\uparrow}\left(e^{-\lambda_0 T_0 - \theta_1 M_1}\cdots e^{-\lambda_k T_k - \theta_{k+1}M_{k+1}}\right)$$
$$= \mathbb{E}_a^p\left(e^{-\lambda_0 T_0 - \theta_1 M_1}\cdots e^{-\lambda_{k-1}T_{k-1} - \theta_k M_k}\frac{M_k}{a}\left(\mathbb{E}_1^p\left(e^{-\lambda_k T_0' - \theta_{k+1}M_1'}\right)\right)^{M_k - 1}\mathbb{E}_1^p\left(e^{-\lambda_k T_0' - \theta_{k+1}M_1'}M_1'\right)\right).$$

Now using the Markov property of $\{(T_n, M_{n+1}), n \in \mathbb{Z}_+\}$ and writing the terms suitably, we get

$$\mathbb{E}_a^{p\uparrow}\left(e^{-\lambda_0 T_0 - \theta_1 M_1}\cdots e^{-\lambda_k T_k - \theta_{k+1}M_{k+1}}\right)$$
$$= e^{-\frac{1}{n}\log\mathbb{E}_n^p\left(e^{-\lambda_k T_0' - \theta_{k+1}M_1'}\right)}\mathbb{E}_1^p\left(e^{-\lambda_k T_0' - \theta_{k+1}M_1'}M_1'\right)$$
$$\times \mathbb{E}_a^p\left(e^{-\lambda_0 T_0 - \theta_1 M_1}\cdots e^{-\lambda_{k-1}T_{k-1}}e^{-\left(\theta_k - b_2\log\mathbb{E}_{b_2^{-1}}^p\left(e^{-\lambda_k T_0' - \theta_{k+1}M_1'}\right)\right)M_k}\frac{M_k}{a}\right).$$

Taking in the previous equality $(\lambda_{i-1}, \theta_i) = (b_1 s_{i-1}, b_2 t_i)$, $i = 1, ..., k+1$ and $(a, p) = (a(n), p(n))$, we use hypothesis (1.13) and (1.14) to obtain

$$\mathbb{E}_{a(n)}^{p(n)\uparrow} \left(\prod_{i=1}^{k+1} e^{-b_1 s_{i-1} T_{i-1} - b_2 t_i M_i} \right)$$

$$\sim e^{-b_2 \log \mathbb{E}_{b_2^{-1}}^{p(n)} \left(e^{-b_1 s_k T_0' - b_2 t_{k+1} M_1'} \right)} \mathbb{E}_1^{p(n)} \left(e^{-b_1 s_k T_0' - b_2 t_{k+1} M_1'} M_1' \right)$$

$$\times \mathbb{E}_{a(n)}^{p(n)\uparrow} \left(e^{-b_1 s_0 T_0 - b_2 t_1 M_1} \cdots e^{-b_1 s_{k-1} T_{k-1} - b_2 (t_k + \kappa(s_k + c t_{k+1})) M_k} \right).$$

Now we have to calculate the limit of each factor. The first one converges towards to 1 thanks to (4.28). Besides, to get

$$\mathbb{E}_1^{p(n)} \left(e^{-b_1 s_k T_0' - b_2 t_{k+1} M_1'} M_1' \right) \xrightarrow[n \to \infty]{} \iota(\kappa_0(s_k + c_\vartheta t_{k+1})), \tag{4.30}$$

we use the convergence (4.29) together the convention $\kappa_0(s) = s$. As in Proposition 4.5, in order to conclude we use the induction hypothesis with $s_{i-1}' + t_i'$, $1 \le i \le k$ defined in (4.23).

A   Proof of Lemma 4.1

A consequence of (1.9) is that the measure defined on $[0, \infty)$ by

$$\mu(x) := \int_0^x z \bar{\pi}^+(z) dz, \quad x \ge 0. \tag{A.1}$$

is such that $x \mapsto \mu(x)$ is $RV_\infty^{2-\alpha}$. Then from the Tauberian-Abelian Theorem (see [4]), its Laplace transform $\mathcal{L}_\mu \in RV_0^{-(2-\alpha)}$ and

$$\mu(x) \sim \frac{1}{\Gamma(3-\alpha)} \mathcal{L}_\mu(1/x), \quad x \to \infty.$$

Observe that

$$\lambda^2 \mathcal{L}_\mu(\lambda) = \int_0^\infty (1 - e^{-\lambda y} - \lambda y) \pi^+(dy) + \int_0^\infty \lambda z (1 - e^{-\lambda z}) \pi^+(dz).$$

Then

$$\lambda^2 \mathcal{L}_\mu(\lambda) = \mathbb{E} \left(1 - e^{-\lambda \xi^+} - \lambda \xi^+ e^{-\lambda \xi^+} \right), \quad \lambda \to 0. \tag{A.2}$$

As consequence of the definition of the measure μ and the approximations above,

$$\bar{\pi}^+(1/\lambda) \sim c_\alpha \mathbb{E} \left(1 - e^{-\lambda \xi^+} - \lambda \xi^+ e^{-\lambda \xi^+} \right), \quad \lambda \to 0, \tag{A.3}$$

where $c_\alpha = 1/\Gamma(3-\alpha)$. Hence for all $x > 0$,

$$\frac{1}{\mathbb{E} \left(1 - e^{-\lambda \xi^+} - \lambda \xi^+ e^{-\lambda \xi^+} \right)} \bar{\pi}^+(x/\lambda) \sim \frac{\bar{\pi}^+(x/\lambda)}{c_\alpha \bar{\pi}^+(1/\lambda)} \xrightarrow[\lambda \to 0]{} c_\alpha x^{-\alpha}. \tag{A.4}$$

We set $r(n) = \left(\mathbb{E} \left[1 - e^{-\xi^+/n} - \xi^+ e^{-\xi^+/n}/n \right] \right)^{-1}$ and define the measure on $(0, \infty)$, $m_n(dy) = r(n)\bar{\pi}^+(ndy)$. The convergence in (A.4) implies

$$m_n(x, \infty) \xrightarrow[n \to \infty]{} \int_x^\infty \frac{c_\alpha}{\alpha} \frac{dy}{y^{1+\alpha}}, \quad \text{for all } x \geq 0.$$

Therefore, for all $0 < x \leq y \leq \infty$

$$m_n(x, y] \xrightarrow[n \to \infty]{} \int_x^y \frac{c_\alpha}{\alpha} \frac{dz}{z^{1+\alpha}}.$$

This implies that the measure on $(0, \infty)$ defined by $m_n(dy) = r(n)\bar{\pi}^+(ndy)$ converges vaguely towards $c_\alpha \frac{dy}{y^{1+\alpha}}$. We also have

$$\int y^2 \mathbb{1}_{\{y \leq x\}} r(n)\pi^+(ndy) \xrightarrow[n \to \infty]{} c_\alpha \int y^2 \mathbb{1}_{\{y \leq x\}} \frac{dy}{y^{1+\alpha}}.$$

Using an argument of monote class to deduce the above convergence over intervals $I \subset (0, \infty)$. Thus we obtain the convergence of the Laplace transform of the measure μ. This complete the proof because μ is regularly varying at infinity with indice $2 - \alpha$ and its Laplace transform satisfies the identity (A.2).

B   Proof of Proposition 4.8

Before proving Proposition 4.8, we would like to present some basic aspect of functional convergence of stochastic process, further details can be found in [8]. Is well know that the law of a Lévy process $\{X_t : t \geq 0\}$ on $\mathbb{R}^d$ is determined by that of random variable X_1, which is infinitely divisible random variable, and according to the Lévy-Khintchine formula has characteristic exponent

$$\Psi(\mathbf{u}) = i\mathbf{u} \cdot \mathbf{b} - \frac{1}{2}\mathbf{u} \cdot c\mathbf{u}^T + \int \left(e^{i\mathbf{u}\cdot\mathbf{x}} - 1 - i\mathbf{u} \cdot \mathbf{h}(\mathbf{x}) \right) F(d\mathbf{x})$$

where $\mathbf{b} \in \mathbb{R}^d$, c is a $d \times d$ symmetric nonnegative matrix, F is a positive measure on $\mathbb{R}^d$ with $F(\{\mathbf{0}\}) = 0$ and $\int (1 \wedge |\mathbf{x}|^2) F(d\mathbf{x}) < \infty$, $\mathbf{h}$ is a truncation function from $\mathbb{R}^d$ to $\mathbb{R}^d$, that is, bounded measurable satisfying

$$\mathbf{h}(\mathbf{x}) = o(|\mathbf{x}|), \quad |\mathbf{x}| \to 0.$$

Hence an infinitely divisible distribution, and therefore a Lévy process, is uniquely characterized by the triple $(\mathbf{b}, c, F)$. Another useful related equality is a $d \times d$ symmetric nonnegative matrix, called the modified second characteristic, and defined as follows

$$\tilde{c}^{ij} = c^{ij} + \int h^i(\mathbf{x})h^j(\mathbf{x})F(d\mathbf{x}), \quad i, j = 1, 2, ..., d.$$

According to Theorem VII.2.9 of [8], if $\{F_n\}_{n \geq 1}$ is a sequence of infinitely divisible distributions on $\mathbb{R}^d$. Then $F_n \to F$ weakly if and only if

$$\mathbf{b}_n \to \mathbf{b}$$
$$\tilde{c}_n \to \tilde{c}$$
$$F_n(\mathbf{g}) \to F(\mathbf{g}) \quad \text{for all } \mathbf{g} \in C_1(\mathbb{R}^d),$$

27

where $C_1(\mathbb{R}^d)$ is a convergence-determining class for the weak convergence induced by all continuous bounded non-negative functions $\mathbb{R}^d \to \mathbb{R}$, vanishing at the origin and with limit at infinity. We stress that here we consider characteristics relatives to a continuous truncation function, $\mathbf{h}$.

In a more general sense, a d-dimensional semimartingale W, has associated a characteristic triplet (B, C, ν) consisting in:

- $B = (B^i)_{i \leq d}$ a predictable process with components of finite variation over each interval $[0, t]$.

- $C = (C^{ij})_{i,j \leq d}$ a continuous process, namely

$$C^{ij} = \langle W^{i,c}, W^{j,c} \rangle,$$

  where W^c is the continuous martingale part of W.

- ν a predictable random measure on $\mathbb{R}_+ \times \mathbb{R}^d$.

A second modified characteristic $\tilde{C}$ is also defined,

$$\tilde{C}_t^{ij} = C_t^{ij} + (h^i h^j) * \nu_t - \sum_{s \leq t} \left(\int h^i(\mathbf{x}) \nu(\{s\} \times d\mathbf{x}) \right) \left(\int h^j(\mathbf{w}) \nu(\{s\} \times d\mathbf{w}) \right).$$

If W has no fixed times of discontinuity, in which case B is continuous, and $|h(x)|^2 * \nu_t < \infty$, it reduces to

$$\tilde{C}_t^{ij} = C_t^{ij} + (h^i h^j) * \nu_t.$$

According to Theorem VII.3.4 of [8], the necessary and sufficient conditions to assure the functional convergence of a sequence of semimartingales W^n towards W are given also in terms of their characteristics:

$$\sup_{s \leq t} |B_s^n - B_s| \to 0, \text{ for all } t \geq 0,$$
$$\tilde{C}^n \to \tilde{C}, \text{ for all } t \in D,$$
$$\mathbf{g} * \nu_t^n \to \mathbf{g} * \nu_t, \text{ for all } t \in D, \mathbf{g} \in C_1(\mathbb{R}^d), \tag{B.1}$$

where D is a dense subset of $\mathbb{R}_+$.

We now turn to prove the convergence claimed in Proposition 4.8. In order to apply the Theorem VII.3.4 of [8], first we will prove the convergence of the characteristics of the process

$$\tilde{\mathbf{S}}_{N_{r(n)t}}^n - (r(n)t/n, 0), \quad t \geq 0, \tag{B.2}$$

where

$$\tilde{\mathbf{S}}_k^n = \sum_{i=1}^k \left(\xi_i^{(cn)}/n, \xi_i^{(mn)}/r(n)p(n) \right), \quad k \in \mathbb{N},$$

and $\{N_t, t \geq 0\}$ is a Poisson process with parameter one, independent of the sequence

$$\xi^{(n)} = \left\{ \left(\xi_k^{(cn)}, \xi_k^{(mn)} \right) : k \in \mathbb{Z}_+ \right\}.$$

The following Lemma establishes the previous statement. We will use this result as a device to study the characteristics of $\bar{\mathbf{S}}_{\lfloor r(n)t \rfloor}^n$, which are closely related to those of $\tilde{\mathbf{S}}_{N_{r(n)t}}^n - (r(n)t/n, 0)$.

28

Lemma B.1. *The process defined in (B.2) is a semimartingale with characteristics relatives to a continuous truncation function* $\mathbf{h}$ *given by*

$$
\begin{aligned}
\mathbf{b}_t^n &= r(n)t\mathbb{E}\left[\mathbf{h}\left(b(n)\xi^{(n)}\right)\right] - (r(n)t/n, 0), \\
c_t^{n,ij} &= 0, \quad \tilde{c}_t^{n,ij} = r(n)t\mathbb{E}\left[h_i\left(b(n)\xi^{(n)}\right)h_j\left(b(n)\xi^{(n)}\right)\right], \quad i,j = 1,2, \\
F_t^n(d\mathbf{x}) &= r(n)t\pi(d\mathbf{x}),
\end{aligned}
\tag{B.3}
$$

where $b(n)\xi^{(n)} = \left(\xi^{(cn)}/n, \xi^{(mn)}/r(n)p(n)\right)$, $\pi^{(n)}(d\mathbf{x}) = \mathbb{P}(\xi^{(cn)} \in dx_1, \xi^{(mn)} \in dx_2)$. *Moreover, in the regime (1.9) and (1.14) we have the following weak convergence in the sense of Skorohod topology*

$$
\left(\left(\tilde{\mathbf{S}}_{N_{r(n)t}}^n - (r(n)t/n, 0) : t \geq 0\right), \mathbb{P}_{a(n)}^{p(n)}\right) \Longrightarrow (X_t, t \geq 0),
\tag{B.4}
$$

where X_t *is a spectrally positive* α*-stable process with parameter* $\alpha \in (1,2)$. *In particular, we obtain the convergence of the characteristics in (B.3) towards those relatives to* $((X_t, t) : t \geq 0)$ *and characteristic exponent* $c_\alpha|\lambda|^\alpha$, *that is*

$$
\begin{aligned}
\mathbf{b}_t &= \left(t\left(\int_{(0,\infty)}\lambda(h(y) - y)c_\alpha y^{-(\alpha+1)}dy\right), t\right), \\
c_t^{ij} &= 0, \quad \tilde{c}_t^{ij} = \mathbb{E}\left[h_i\left(X_t\right)h_j\left(X_t\right)\right] \quad i,j = 1,2, \\
F_t(d\mathbf{x}) &= tc_\alpha x_1^{-(\alpha+1)}dx_1\delta_0(dx_2),
\end{aligned}
\tag{B.5}
$$

where h *is a continuous truncation function from* $\mathbb{R}$ *to* $\mathbb{R}$.

Proof of Lemma B.1. Note that for $\mathbf{u} = (\lambda, \theta) \in \mathbb{R}^2$,

$$
\mathbb{E}\left(e^{i\mathbf{u}\cdot\tilde{\mathbf{S}}_{N_{r(n)t}}^n}\right) = e^{r(n)t\left(\psi_n\left(\frac{\lambda}{n}, \frac{\theta}{r(n)p(n)}\right) - 1\right)}, \quad t \geq 0.
\tag{B.6}
$$

Then the exponent in the righthand side of the previous equality can be written as follows

$$
t\left[i\int_{(0,\infty)}\int_{(0,\infty)}\mathbf{u}\cdot\mathbf{h}(b(n)\mathbf{x})r(n)\pi(d\mathbf{x}) + \int_{(0,\infty)}\int_{(0,\infty)}\left(e^{i\mathbf{u}\cdot b(n)\mathbf{x}} - 1 - i\mathbf{u}\cdot\mathbf{h}(b(n)\mathbf{x})\right)r(n)\pi^{(n)}(d\mathbf{x})\right],
\tag{B.7}
$$

where $b(n)\mathbf{x} = (x_1/n, x_2/r(n)p(n))$, $\pi^{(n)}(d\mathbf{x}) = \mathbb{P}(\xi^{(cn)} \in dx_1, \xi^{(mn)} \in dx_2)$. From here $\tilde{\mathbf{S}}_{N_{r(n)t}}$ is infinitely divisible, also we can deduce that the characteristics of the process $\{\tilde{\mathbf{S}}_{N_{r(n)t}}^n - (r(n)t/n, 0) : t \geq 0\}$ are given by (B.3). Thanks to Theorem II.3.11 of [8] this process is a Lévy process and a semimartingale.

Besides, to get the convergence in (B.4) we shall prove the convergence of the characteristic functions. This fact is verified using Corollary 4.3, together with the fact that conditionally to $\xi^{(+)} = k$ the distribution of $\xi^{(cn)}$ is Binomial with parameter $(k, 1-p(n))$, as well as the assumption that ξ^+ has mean 1.

$$
t\int_{(0,\infty)}\left(e^{-((1-p(n))(1-e^{i\lambda/n})-p(n)(1-e^{i\theta/r(n)p(n)}))y} - 1 - i\mathbf{u}\cdot\mathbf{h}\left(\frac{y}{n}(1-p(n)), \frac{y}{r(n)p(n)}p(n)\right)\right)r(n)\pi^+(dy)
$$

$$
+ ti\int_{(0,\infty)}\left(\mathbf{u}\cdot\mathbf{h}\left(\frac{y}{n}(1-p(n)), \frac{y}{r(n)p(n)}p(n)\right) - \left(\frac{\lambda}{n}(1-p(n)) + \frac{\theta}{r(n)p(n)}p(n)\right)y\right)r(n)\pi^+(dy)
$$

$$
+ i\left(\frac{\lambda}{n}(1-p(n)) + \frac{\theta}{r(n)p(n)}p(n)\right)r(n)t.
$$

29

Then by the continuity of $\mathbf{h}$ and the assumptions (1.13) and (1.10), the previous display becomes

$$\sim t \int_{(0,\infty)} \left(e^{-((1-p(n))(1-e^{\mathrm{i}\lambda/n})-p(n)(1-e^{\mathrm{i}\theta/r(n)p(n)}))y} - 1 - \mathrm{i}\lambda h\left(\frac{y}{n}\right) \right) r(n)\pi^+(dy)$$

$$+ t\mathrm{i} \int_{(0,\infty)} \left(\lambda h\left(\frac{y}{n}\right) - \left(\frac{\lambda}{n}(1-p(n)) + \frac{\theta}{r(n)p(n)}p(n)\right)y \right) r(n)\pi^+(dy)$$

$$+ \mathrm{i}\left(\frac{\lambda}{n}(1-p(n)) + \frac{\theta}{r(n)p(n)}p(n)\right) r(n)t.$$

where h is the truncation function from $\mathbb{R}$ to $\mathbb{R}$ obtained as projection of $\mathbf{h}$ in the second coordinate, that is,

$$\sim t \int_{(0,\infty)} \left(e^{-((1-p(n))(1-e^{\mathrm{i}\lambda/n})-p(n)(1-e^{\mathrm{i}\theta/r(n)p(n)}))ny} - 1 - \mathrm{i}\lambda h(y) \right) r(n)\pi^+(ndy)$$

$$+ t\mathrm{i} \int_{(0,\infty)} \left(\lambda h(y) - \left(\lambda(1-p(n)) + \frac{\theta}{r(n)p(n)}np(n)\right)y \right) r(n)\pi^+(ndy)$$

$$+ \mathrm{i}\left(\frac{\lambda}{n}(1-p(n)) + \frac{\theta}{r(n)p(n)}p(n)\right) r(n)t. \tag{B.8}$$

Finally we have the convergence

$$\mathbb{E}\left(e^{\mathrm{i}\mathbf{u}\cdot(\tilde{\mathbf{S}}^n_{N_{r(n)t}} - (r(n)t/n,0))} \right) \xrightarrow[n\to\infty]{} e^{t\left(\int_{(0,\infty)}\left(e^{\mathrm{i}\lambda y}-1-\mathrm{i}\lambda y\right)c_\alpha y^{-(\alpha+1)}dy\right)+\mathrm{i}t\theta}, \tag{B.9}$$

where c_α is a constant depending on α that appears in Lemma 4.1. Indeed, this is a consequence of the arguments in this Lemma together with the hypothesis (1.13)

$$\int_{(0,\infty)} \left(e^{-((1-p(n))(1-e^{\mathrm{i}\lambda/n})-p(n)(1-e^{\mathrm{i}\theta/r(n)p(n)}))ny} - 1 - \mathrm{i}\lambda h(y) \right) r(n)\pi^+(ndy)$$

$$\xrightarrow[n\to\infty]{} t \int_{(0,\infty)} \left(e^{\mathrm{i}\lambda y} - 1 - \mathrm{i}\lambda h(y) \right) c_\alpha y^{-(\alpha+1)}dy,$$

while the second adding converges to

$$\mathrm{i}t \int_{(0,\infty)} \lambda(h(y) - y)c_\alpha y^{-(\alpha+1)}dy.$$

To end we observe that $r(n) \in RV_\infty^\alpha$ with $\alpha \in (1,2]$ and $p(n) \sim cn^{-1}$ implies that

$$\frac{r(n)p(n)}{n} \xrightarrow[n\to\infty]{} 0.$$

Then

$$\mathrm{i}\left(\frac{\lambda}{n}(1-p(n)) + \frac{\theta}{r(n)p(n)}p(n)\right) r(n)t - \mathrm{i}\frac{\lambda}{n}r(n)t,$$

converges to $\mathrm{i}\theta t$. From (B.9) the characteristic of (X_t, t) are given by (B.5). As a consequence of Theorem VII.2.9 of [8] we have the convergence of the characteristics. Finally the characteristic exponent of X_t is $c_\alpha\lambda^\alpha$. ∎

We have now all the elements to prove Proposition 4.8.

Proof of Proposition 4.8. Thanks to Theorem 2.3.11 of [8], $\tilde{\mathbf{S}}^{(n)}_{\lfloor r(n)t \rfloor}$ is a semimartingale with characteristics relatives to $\mathbf{h} : \mathbb{R}^2 \to \mathbb{R}^2$, given by

$$B^n_t = \lfloor r(n)t \rfloor \mathbb{E}\left[\mathbf{h}\left(b(n)\xi^{(n)}\right)\right] - \left(\lfloor r(n)t \rfloor / n, 0\right)$$
$$C^{n,ij}_t = 0, \quad \tilde{C}^{n,ij}_t = \lfloor r(n)t \rfloor \left(\mathbb{E}\left[h_i\left(b(n)\xi^{(n)}\right)h_j\left(b(n)\xi^{(n)}\right)\right] - \mathbb{E}\left[h_i\left(b(n)\xi^{(n)}\right)\right]\mathbb{E}\left[h_j(b(n)\xi^{(n)})\right]\right) \quad i,j = 1,2.$$
$$\mathbf{g} * \nu^n_t = \lfloor r(n)t \rfloor \mathbb{E}\left(\mathbf{g}\left(b(n)\xi^{(n)}\right)\right) \quad \text{for } \mathbf{g} \geq 0 \text{ Borel,}$$

Now only remains to verify the conditions (B.1) where the limit characteristics agree with these in (B.5). In this direction we recall that $\mathbf{b}^n_t = r(n)t\mathbb{E}\left[\mathbf{h}\left(b(n)\xi^{(n)}\right)\right] - (r(n)t/n, 0)$ and observe

$$|B^n_s - \mathbf{b}_s| \leq |\lfloor r(n)s \rfloor - r(n)s|\, \mathbb{E}\left[\mathbf{h}\left(b(n)\xi^{(n)}\right)\right] + |(r(n)s/n, 0) - (\lfloor r(n)s \rfloor/n, 0)| + |\mathbf{b}^n_s - \mathbf{b}_s|.$$

Then using the properties of the floor function we obtain

$$|B^n_s - \mathbf{b}_s| \leq \mathbb{E}\left[\mathbf{h}\left(b(n)\xi^{(n)}\right)\right] + |(s/n, 0)| + |\mathbf{b}^n_s - \mathbf{b}_s|.$$

Thus by the convergence of $\mathbf{b}^n_t$ established in the previous lemma, together with $r(n) \to \infty$

$$\sup_{s \leq t} |B^n_s - \mathbf{b}_s| \leq \mathbb{E}\left[\mathbf{h}\left(b(n)\xi^{(n)}\right)\right] + |(t/n, 0)| + |\mathbf{b}^n_t - \mathbf{b}_t| \xrightarrow[n \to \infty]{} 0,$$

hence we have the first condition in (B.1). In order to determine the second one, let $b^{n,i}_t$ be the i-th coordinate of $\mathbf{b}^n_t$, $i = 1, 2$. Once more, applying the properties of the floor function we have

$$\left(1 - \frac{1}{r(n)t}\right)\tilde{c}^{n,ij}_t - \frac{1}{r(n)t}b^{n,i}_t b^{n,j}_t - \frac{r(n)t}{n}\mathbb{E}\left[h_2\left(b(n)\xi^{(n)}\right)\right]$$
$$\leq \tilde{C}^{n,ij}_t \leq \tilde{c}^{n,ij}_t + \frac{1 - r(n)t}{(r(n)t)^2}b^{n,i}_t b^{n,j}_t + \frac{1 - r(n)t}{n}\mathbb{E}\left[h_2\left(b(n)\xi^{(n)}\right)\right].$$

Also
$$\left(1 - \frac{1}{r(n)t}\right)\tilde{c}^{n,22}_t - \frac{1}{r(n)t}(b^{n,2}_t)^2 \leq \tilde{C}^{n,22}_t \leq \tilde{c}^{n,22}_t + \frac{1 - r(n)t}{(r(n)t)^2}(b^{n,2}_t)^2,$$

and
$$\left(1 - \frac{1}{r(n)t}\right)\tilde{c}^{n,11}_t - \frac{1}{r(n)t}(b^{n,1}_t)^2 \leq \tilde{C}^{n,11}_t \leq \tilde{c}^{n,11}_t + \frac{1 - r(n)t}{(r(n)t)^2}(b^{n,1}_t)^2 + b^{n,1}_t + \frac{1}{n^2}[r(n) - 1].$$

As consequence of the convergence $\mathbf{b}^n_t \to \mathbf{b}_t$, $b^{n,i}_t$ converges for $i = 1, 2$. Then the above inequalities impliy $\tilde{C}^{n,ij}_t \to \tilde{c}^{ij}_t$, because $r(n) \in RV^\alpha_\infty$ and $\alpha \in (1, 2)$. It is easily prooved that also $\mathbf{g}*\nu^n_t \to \mathbf{g}*F_t$ for all g using that $F^n_t = r(n)t\pi(d\mathbf{x})$ converges to F_t, as we proved in the previous lemma, together with properties of the floor function. ∎

References

[1] K. B. Athreya and P. E. Ney. *Branching processes*. Springer-Verlag, New York-Heidelberg, 1972. Die Grundlehren der mathematischen Wissenschaften, Band 196.

[2] J. Bertoin. *Lévy processes*, volume 121 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 1996.

[3] J. Bertoin. A limit theorem for trees of alleles in branching processes with rare neutral mutations. *Stochastic Process. Appl.*, 120(5):678–697, 2010.

[4] N. H. Bingham, C. M. Goldie, and J. L. Teugels. *Regular variation*, volume 27 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1987.

[5] B. Chauvin. Arbres et processus de Bellman-Harris. *Ann. Inst. H. Poincaré Probab. Statist.*, 22(2):209–232, 1986.

[6] K. L. Chung and J. B. Walsh. *Markov processes, Brownian motion, and time symmetry*, volume 249 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer, New York, second edition, 2005.

[7] W. Feller. *An introduction to probability theory and its applications. Vol. II.* Second edition. John Wiley & Sons, Inc., New York-London-Sydney, 1971.

[8] J. Jacod and A. N. Shiryaev. *Limit theorems for stochastic processes*, volume 288 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1987.

[9] O. Kallenberg. *Foundations of modern probability*. Probability and its Applications (New York). Springer-Verlag, New York, second edition, 2002.

[10] K. Kawazu and S. Watanabe. Branching processes with immigration and related limit theorems. *Teor. Verojatnost. i Primenen.*, 16:34–51, 1971.

[11] A. Lambert. Quasi-stationary distributions and the continuous-state branching process conditioned to be never extinct. *Electron. J. Probab.*, 12:no. 14, 420–446, 2007.

[12] A. Lambert. Population dynamics and random genealogies. *Stoch. Models*, 24(suppl. 1):45–163, 2008.

[13] Z. Li. *Measure-valued branching Markov processes*. Probability and its Applications (New York). Springer, Heidelberg, 2011.

[14] W. Whitt. *Stochastic-process limits*. Springer Series in Operations Research. Springer-Verlag, New York, 2002. An introduction to stochastic-process limits and their application to queues.